

DOCUMENT RESUME

ED 429 133

TM 029 742

AUTHOR August, Diane; Muraskin, Lana
TITLE Strengthening the Standards: Recommendations for OERI Peer Review. Summary Report. Draft, January 30, 1999.
INSTITUTION August and Associates, Bethesda, MD.
SPONS AGENCY National Educational Research Policy and Priorities Board (ED/OERI), Washington, DC.
REPORT NO PPB-1999-6314g
PUB DATE 1999-01-30
NOTE 130p.; "Prepared under the auspices of a panel of experts: Christopher T. Cross, Carl F. Kaestle, Sharon Lewis, Penelope Peterson, and Judith Sunley."
PUB TYPE Reports - Descriptive (141)
EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS *Educational Research; Elementary Secondary Education; *Evaluation Methods; *Grants; Higher Education; *Peer Evaluation; Program Evaluation; *Research Methodology; *Standards
IDENTIFIERS *Office of Educational Research and Improvement

ABSTRACT

Under a legislative mandate, the National Educational Research Policy and Priorities Board has been charged with the development of standards for research conducted under the auspices of the Office of Educational Research and Improvement. The first standards developed are those for peer review. This study is a retrospective review of the implementation of the standards for evaluation and peer review of grant and cooperative agreement programs. The review focuses on two years of competitions (fiscal years 1996 and 1997) and two types of competitions for field-initiated studies (FIS) and for research centers. The study includes 20 randomly selected FIS panels and all center competitions. Findings from this review lead to several recommendations. The first is that standing panels of reviewers should be established in each institute. These reviewers should serve staggered terms and be the core for each review. Options are defined for the FIS process and the scoring of applications that have not been rejected or deferred. Other recommendations for the peer review process include improved professional development and improved feedback to unsuccessful applicants. Three appendixes contain the standards, a technical review form, and evaluation criteria for peer review. (SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

TM

Prepublication Copy

STRENGTHENING THE STANDARDS: RECOMMENDATIONS FOR OERI PEER REVIEW

**Summary Report
Draft January 30, 1999**

August and Associates

TM029742

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it

☐ Minor changes have been made to
improve reproduction quality

• Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy

BEST COPY AVAILABLE

STRENGTHENING THE STANDARDS: RECOMMENDATIONS FOR OERI PEER REVIEW

**Summary Report
Draft January 30, 1999**

*U.S. Department of Education
National Educational Research Policy and Priorities Board*

*Prepared Under the Auspices of a Panel of Experts:
Christopher T. Cross, Carl F. Kaestle, Sharon Lewis,
Penelope Peterson, and Judith Sunley*

August and Associates
Diane August
Bethesda, MD

Lana D. Muraskin
Washington, D.C.

ACKNOWLEDGMENTS

Many people contributed to this report during the writing and review process. The authors are particularly indebted to the expert panel that provided guidance and thoughtful reviews over the course of the project. Panelists include Christopher T. Cross, Carl F. Kaestle, Sharon Lewis, Penelope Peterson, and Judith Sunley. The authors especially appreciate the help of the government COTR, Kathy FitzGerald, who tracked down materials, provided ongoing information, and reviewed multiple drafts of the report. Diane Pelavin and Cecily Darden from the American Institutes for Research, the prime contractor for the report, provided logistical support. Carole Spalding provided administrative assistance and Rona Briere editorial assistance.

Many other U.S. Department of Education officials also provided information about the peer review process and helpful comments on multiple drafts. They include the OERI Institute Directors—Joseph Conaty, Edward Fuentes, Deborah Inman (formerly Institute Director for the Policy Institute), Naomi Karp, and Carole Lacampagne. Other Department staff who contributed to the report include Judith Anderson, Clara Lawson-Holmes, Jacqueline Jenkins, Beth Fine, Karen Suagee, Gilbert Garcia, Veda Bright, Seresa Simpson, Duc-Le To, Delores Monroe, Jerome Lord, James Fox, Ron Hanson, Patricia Knight, Margo Anderson, Cynthia Stewart, Elizabeth Payer, Brian LeKander, and Doris Andres.

Eve Bither and Thelma Leenhouts from the National Educational Research Policy and Priorities Board (NERPPB) attended panel meetings and provided valuable comments about the peer review process. NERPPB members provided oversight of our work at critical junctures.

We are especially grateful to John Bruer, Kenji Hakuta, Glenda Lappan, Bob Marley, Joyce Muhlestein, Claire Pelton, and Alba Ortiz who make up the NERPPB subgroup on peer review.

We thank Jerry Sroufe, Executive Director of the American Education Research Association, and David Johnson, Executive Director of the Federation of Behavioral, Psychological, and Cognitive Sciences, who provided information about the peer review process at OERI and other agencies. We also appreciate the help of staff from other agencies and organizations who provided information about peer review; they include Candace Katz and Ken Colson from the National Endowment for the Humanities, Janice Earle from the National Science Foundation, and Julie Jessen from the Spencer Foundation.

EXECUTIVE SUMMARY

The literature on the uses of peer review sets forth several commonly cited goals of the peer review process that provide a context for the present study. They include efficacy—ensuring that scientifically appropriate research is supported; efficiency—ensuring that the research selected is cost-effective; fairness—ensuring that bias is removed from the application review process; and public accountability—ensuring that the public will is reflected and the public interest protected in the selection of grantees.

On March 31, 1994, President Clinton signed Public Law 103-227, which includes Title IX, the Educational Research, Development, and Improvement Act of 1994. This legislation mandated that the Assistant Secretary of the Office of Educational Research and Improvement (OERI), in consultation with the newly established National Educational Research Policy and Priorities Board (NERPPB), develop standards to govern the conduct and evaluation of all research, development, and dissemination activities carried out by the Office, and to ensure that such activities meet the highest standards of professional excellence. These standards were to be developed in three phases, the first of which would address OERI's process for peer review. According to the legislation, peer review standards would, at a minimum (1) describe the general procedures to be used by each peer review panel in its operations; (2) describe the procedures to be used in evaluating applications for grants, cooperative agreements, and contracts; and (3) specify the criteria and factors to be considered in making such evaluations.

In accordance with these legislative requirements, the Assistant Secretary established final regulations to set standards for the evaluation of applications for grants and cooperative agreements and proposals for contracts, to take effect October 16, 1995. These standards were developed by the Assistant Secretary, in consultation with NERPPB. They were first published in a notice of proposed rulemaking in the *Federal Register*, June 7, 1995, to invite comments from interested parties. NERPPB gave final approval for these standards in fiscal year (FY) 1996 after all public comments had been received, discussed, and addressed as appropriate (Code of Federal Regulations [CFR], Title 34, Part 700). To date these standards for peer review have been applied in 2 years of competitions.

This study is a retrospective review of the implementation of the Office of Educational Research and Improvement (OERI)/National Educational Research Policy and Priorities Board (NERPPB) standards for evaluation and peer review of grant and cooperative agreement applications. The review focuses on 2 years of competitions—fiscal years 1996 and 1997—and two types of competitions—for field-initiated studies (FIS) and for research centers. The study includes 20 randomly selected FIS panels and all center competitions. The review was carried out by OERI under contract, and was overseen by an expert panel that helped frame the study, direct its progress, and craft its recommendations.¹ These recommendations are designed to assist OERI and NERPPB in considering whether to make changes in the standards

¹ Panel members included: Christopher Cross, President of the Council for Basic Education; Carl F. Kaestle, Professor of Education, History, and Public Policy, Brown University; Sharon Lewis, Director of Research, Council of Great City Schools; Penelope Peterson, Dean of the School of Education and Social Policy, Northwestern University; and Judith Sunley, Assistant to the Director for Social Policy and Planning, the National Science Foundation.

or their application. This summary highlights the study recommendations that are most central to improving the peer review process.

First, standing panels of 25 to 30 reviewers should be established in each institute. Reviewers should be carefully selected to ensure that each meets the criteria established by the standards. Panels should be constituted to ensure ethnic, racial, geographic, and gender diversity. Moreover, a balance between senior and junior scholars should be sought to provide professional development for junior scholars and "enliven" the thinking of senior scholars. Proposed panelist slates should be approved by the institute directors and the Assistant Secretary for OERI, with consultation from the NERPPB.

The reviewers on these standing panels should serve set (e.g., staggered 3-year) terms and form the core of reviewers for each institute. For the center competitions, a subset of standing panelists should be used. Decisions about which panelists to select for a center competition and the number needed should be based on the applications received for a particular competition. The subpanelists could also serve as midterm reviewers, thus ensuring consistency in the review process.

For field-initiated studies (FIS) competitions, there are two options for the review process. The first would entail the formation of six- to eight-member subpanels from the membership of the standing panel; these subpanels would provide the first tier of review. The first-tier review process would function much like the current process, except the subpanels would comprise primarily standing panelists and would be expanded from three members to six to eight members to provide a broader context for the review. Applications would be allocated to subpanels on the basis of the panelists' subject area expertise and experience. If the review

of some applications required special technical expertise, the subpanels could be supplemented with ad hoc reviewers. During each review cycle, the team leaders of each subpanel would meet an additional day for a second-tier review to rate all the top-ranked applications from the first-tier subpanels.

A second option for the FIS review process resembles the process used at the National Institutes of Health, where the entire panel reads all applications. At NIH it is typical for a group to review 75 to 100 applications at each meeting. Each member is asked to prepare detailed reviews for a dozen or more applications. The meetings are conducted by a chair who is a peer, assisted by a staff member. Those preparing the written reviews lead the discussion of the applications assigned to them. Each application is discussed and considered. Decisions not to recommend for further consideration are made by majority vote.² If a member disagrees, he or she can submit a minority report, and when there are two or more dissenting members, a minority report must be drafted. Members who cannot assess the merits of a proposal can abstain from voting, although abstentions are not encouraged. Review of applications can also be deferred (perhaps for a site visit or to obtain additional information).

Those applications not rejected or deferred are assigned a priority score by each member. These scores are averaged by the staff member after the meeting. In addition, a summary statement for each application is prepared for transmittal to the council and the applicant by the staff person involved in the review. The statement includes a percentile ranking for the application against a reference base of all applications reviewed by the

² This procedure is not permissible under the current standards.

committee over three meetings, including applications not recommended for funding or deferred. The written comments of panel members and the panel discussions are the basis for these summary statements.

Standing panels might also be involved in other activities, such as recommending how OERI could help foster research in a particular area in which good applications had not been received; helping to select new panelists and ad hoc reviewers; and reviewing grant-produced products, especially once the Phase 3 standards (for review of completed research) have been put in place. Panels might also provide continuity in the assessment of applications so that rejected applications that had been revised and resubmitted would be reviewed by at least some of the same people. In addition, panelists could serve on midterm review teams for existing centers.

Given that some institutes receive up to 200 applications annually, methods for reducing reviewer workload should also be considered. Several possibilities are elaborated in this report. They include, for example, the use of preliminary reviews to reduce the number of full applications receiving a detailed evaluation, and the use of pre-applications, with only a subset of applicants being asked to prepare a full application.

This report also makes other recommendations for improving the OERI peer review process. First, professional development should be enhanced for Department of Education staff, especially those new to the process, as well as for applicants and reviewers. Additionally, reviewers would benefit from questions to guide their reviews and from elaborated scoring rubrics. The standards would benefit from clarification in several areas: the term "multiple perspectives" should be further defined to ensure that panel membership is

balanced by disciplinary background, theoretical orientation, methodological approach, and research role, as well as gender, race/ethnicity, and geographic location; conflict of interest should be defined to include professional relationships as a source of conflict; and priorities should not be listed for FIS competitions. Modifying the review criteria and weightings would also enhance the process: the meaning of "national significance" should be clarified; for center competitions, the project design criterion should be elaborated; and weighting for the management criterion should be increased. Standardization of scores should be eliminated as well; the use of second-tier panels and standing panels would make this process unnecessary. Finally, more detailed feedback should be provided to unsuccessful applicants, and the use of technology in the peer review process should be explored.

CONTENTS

Chapter	Page Number
1 OVERVIEW	CH 1-1
Background for the Study	
Study Design,	
Organization of This Report	
2 THE NATURE OF THE STANDARDS	CH 2-1
Qualifications of Peer Reviewers	
Selection of Peer Reviewers	
Conduct of the Competitions	
Evaluating, Ranking, and Selection	
of Award Recipients	
Procedures Not Specified by the Standards	
3 DESCRIPTION OF THE COMPETITIONS	CH 3-1
Field-Initiated Studies Research Grant Program	
Fiscal Years 1996 and 1997	
National Educational Research and	
Development Center Program Competitions	
4 FINDINGS	CH 4-1
Substantive Fit Between Reviewers and Applications	
Evaluation of Reviews	
Case Studies	

5 STUDY CONCLUSIONS AND RECOMMENDATIONS

CH 5-1

Appropriateness and Usefulness of the Standards
Findings and Recommendations
Summary of Recommendations

RESOURCE MATERIAL

R-1

APPENDICES

A Standards for Peer Review
B Technical Review Form
C Evaluation Criteria for Peer Review

OVERVIEW

Background for the Study

On March 31, 1994, President Clinton signed Public Law 103-227, which includes Title IX, the Educational Research, Development, and Improvement Act of 1994. This legislation mandated that the Assistant Secretary of the Office of Educational Research and Improvement (OERI), in consultation with the newly established National Educational Research Policy and Priorities Board (NERPPB), develop standards to govern the conduct and evaluation of all research, development, and dissemination activities carried out by the Office, and to ensure that such activities meet the highest standards of professional excellence. These standards were to be developed in three phases, the first of which would address OERI's process for peer review. According to the legislation, peer review standards would, at a minimum (1) describe the general procedures to be used by each peer review panel in its operations; (2) describe the procedures to be used in evaluating applications for grants, cooperative agreements, and contracts; and (3) specify the criteria and factors to be considered in making such evaluations.

In accordance with these legislative requirements, the Assistant Secretary established final regulations to set standards for the evaluation of applications for grants and cooperative agreements and proposals for contracts, to take effect October 16, 1995. These standards were developed by the Assistant Secretary, in consultation with NERPPB. They were first

published in a notice of proposed rulemaking in the *Federal Register*, June 7, 1995, to invite comments from interested parties. NERPPB gave final approval for these standards in fiscal year (FY) 1996 after all public comments had been received, discussed, and addressed as appropriate (Code of Federal Regulations [CFR], Title 34, Part 700). To date these standards for peer review have been applied in 2 years of competitions.

NERPPB has asked for a review of the operation of the standards to date. This review is being carried out by OERI under contract, and is being overseen by an expert panel that helped frame the study and directed its progress. The panel members discussed and critiqued multiple versions of the present report and approved its findings and recommendations. The panel includes the following members: Christopher T. Cross, President, Council for Basic Education; Carl F. Kaestle, Professor of Education, History, and Public Policy, Brown University; Sharon Lewis, Director of Research, Council of Great City Schools; Penelope L. Peterson, Dean, School of Education and Social Policy, Northwestern University; and Judith Sunley, Assistant to the Director for Social Policy and Planning, National Science Foundation.

Study Design

This study is a retrospective review of the implementation of the OERI/NERPPB standards for evaluation and peer review of grant and cooperative agreement applications. Specifically, the study charge is to examine (1) whether the standards are appropriate and useful, (2) whether they contribute to fair and high-quality competitions, and (3) how the competitions conducted under the standards have operated and how they may be improved. As appropriate, the study charge includes making recommendations on how to configure and

maintain peer review panels.

The review focuses on 2 years of competitions—fiscal years 1996 and 1997—and two types of competitions—for field-initiated studies (FIS) and for research centers. The unit of analysis is the FIS panel (generally composed of three reviewers and the applications they reviewed) or center competition (composed of five to nine reviewers and the applications they reviewed), rather than applicants or reviewers. The study includes 20 randomly selected FIS panels and all center competitions. The specific research questions, data collection methods, sampling plan, and analysis methods employed are described below.

Research Questions

The specific research questions used for the study are as follows:

- *According to the standards (and other regulations governing OERI peer review), what are the appropriate uses of peer reviewers for FIS and center competitions?*
- *Is the selection process for peer reviewers comprehensive and "unbiased" (e.g., draws reviewers from a wide range of sources and perspectives, selects researchers in various stages of their careers), and does it avoid selection of individuals with conflicts of interest?*
- *Are peer reviewers appropriate to the applications they review? (The standards call for appropriateness with regard to training and experience in the subject area under review, in-depth knowledge of policy and practice in education, and theoretical or methodological approaches in the subject area under review.)*
- *Are peer reviewers adequately instructed or otherwise "trained" for the review process, and is the review process for each type of competition—e.g., mail, in person, one/two stage—carried out effectively?*
- *Does the peer review process yield reviews that provide the information needed to make funding decisions? That is, do the reviews supply the necessary content, including concise written comments, re-reviews after discussion, numerical scores,*

and other recommendations as stipulated in law/regulations?

- *How are funding decisions made? To what extent are peer reviews used in those decisions (e.g., to establish rankings of grant/cooperative agreement applications to provide additional information for decision making)?*
- *What changes/reforms in the peer review system could be recommended to help improve the system?*

Data Collection

To conduct the inquiry outlined above, a number of data collection strategies were undertaken. Interview protocols for applicants and reviewers required Office of Management and Budget (OMB) review. These interviews qualified as customer surveys under guidelines of the Paperwork Reduction Act and received an expedited OMB review. The following are the data collection strategies employed:

- Review of OERI legislation, regulations, standards and other documents related to peer review.
- Conduct of initial open-ended interviews with key federal officials involved in FIS and center competitions to understand the overall review process and use of peer reviewers. Initial interviewees included key staff of the five institutes that have administered the two rounds of FIS competitions and nine center competitions held since the standards were adopted. This initial set of interviews helped us further refine our research questions, identify additional staff for further inquiry, and collect important documents for analysis.
- Semistructured interviews with involved Department of Education (ED) staff on each of the research questions outlined above: appropriate uses of peer reviewers for FIS and center competitions, the peer reviewer selection process, links between the expertise of peer reviewers and that required for proposal reviews, the effectiveness and efficiency of the review process, the quality and usefulness of the reviews generated, the use of reviews in funding decisions, and recommendations for improving the system.
- Semistructured interviews with a sample of reviewers on issues including links

between their expertise and applications reviewed, the quality of instruction and training for the review process, the quality of the overall reviews, the quality and usefulness of review meetings or second-stage reviews in which they participated, and additional observations or recommendations.

- Semistructured interviews with a sample of successful and unsuccessful applicants on issues including the quality of reviews and feedback received and additional observations or recommendations.
- Semistructured interviews with appropriate staff of other Department of Education offices and federal agencies selected for comparison, as well as representatives of educational research associations (e.g., the American Educational Research Association [AERA]), on issues including the appropriateness of OERI law/regulations, peer review practices in other offices or agencies, and recommendations.
- Review of reviews performed on applications of sampled applicants.
- Review of resumes of sampled reviewers.
- A general literature review, with emphasis on the process used by Department of Education offices and other federal agencies selected for comparison.

Sampling Plan

Because we did not have the resources needed to interview all applicants and peer reviewers involved in the FIS and center competitions, we conducted case studies of selected competitions. A total of 367 reviewers reviewed 1,179 applications for the FIS competitions in fiscal years 1996 and 1997. A total of 61 reviewers reviewed 47 applications for the center competitions during those same years.

For the FIS competitions, we studied 20 review panels across the five institutes and the two fiscal years. One panel from each of the five institutes was selected for more in-depth

study because it was identified as successful by the institute director.¹ The remaining 15 panels were selected through a stratified random sample, ensuring that all institutes and both years would be represented. For each of the five "successful" panels, to the extent possible, we interviewed all three panelists, randomly selected two unsuccessful applicants, and interviewed all successful applicants whose applications had been reviewed by that panel.² For the remaining 15 panels we constructed a more limited picture, interviewing two randomly selected panel members and two applicants from each, again to the extent possible. Successful applicants associated with the panel were automatically included and randomly selected from the pool of successful applicants if there were more than two per panel. If there were not two successful applicants, randomly selected unsuccessful applicants were interviewed instead. This approach resulted in 34 interviews of applicants and 40 interviews of reviewers across the competitions. To the extent that they were available, we also reviewed the material associated with the randomly selected applicants and reviewers—38 applications, 100 reviews, and reviewers' resumes.

For the center competitions, we conducted two in-depth case studies—one center competition identified as highly successful by Department of Education staff, and the other identified as problematic. To the extent possible, we interviewed three applicants (including the successful applicant) and three reviewers from each of these competitions. As with the FIS competitions, for each of the other seven center competitions we constructed a more limited picture. To the extent possible, we interviewed two randomly selected reviewers and one

¹ However, only four FIS case studies were conducted. We did not examine one postsecondary FIS panel because one of the authors of this report was a consultant on an application reviewed by that panel.

² We interviewed all the selected reviewers and applicants whom we could locate and who were willing to participate in the study.

randomly selected unsuccessful applicant, as well as the successful applicant. This approach entailed 17 interviews of applicants and 14 interviews of reviewers across the center competitions. To the extent that they were available, we reviewed the materials associated with the randomly selected applicants and reviewers—12 applications, 41 reviews, and reviewers' resumes.

Data Analysis

Analysis of Laws and Regulations. We conducted an analysis of the key laws and regulations that govern OERI grant and cooperative agreement competitions. The analysis considered the requirements of the laws/regulations and their implications for conducting competitions. It also considered any inconsistencies, problematic elements and conflicts among requirements.

Analysis of Interview Data. Because most of the information gathered for the study was based on semistructured interviews, the main analytic approach was qualitative data analysis. Data from interviews with applicants and reviewers was organized into tables by FIS panel and center competition. The data did not reveal enough consistency within panels in the assessment of review quality and reviewer expertise to warrant separate analyses. Thus the data was aggregated across cases and described in narrative form. This approach to analyzing qualitative data is widely used for aggregating and summarizing interview data.

Analysis of the Match Between Applications and Reviewers. To the extent they were available, we collected the resumes of the individuals who reviewed sampled applications, and analyzed the match between the substantive requirements of the applications

and the background, experience, and subject matter expertise of the reviewers.

Analysis of Application Reviews. We developed criteria, based in part on the new standards, for assessing the quality of the written reviews. Measures of quality included thoroughness and detail in comments, discussion of substantive issues in the subject area of the application, and judgments that were amply and expertly justified.

Analysis of Review Process of Six Panels Nominated by OERI Staff. In addition to our overall picture of the competitions, we examined the review process of six panels in greater depth. This examination focused on four fiscal year 1997 FIS panels nominated by institute staff as particularly successful and two center competition panels—one nominated as successful and one as problematic.

Department of Education Staff Interviews. Department of Education staff who were involved in administering and overseeing the competitions and staff from several professional associations were interviewed. Relevant comments and insights have been incorporated in this analysis.³

Review of Peer Review in Other Offices and Agencies. We also examined the peer review process conducted in other offices within the Department of Education; federal agencies; and a private foundation, the Spencer Foundation, that primarily funds education research. The other offices within the Department of Education included the Office of Reform Assistance and Dissemination (specifically the Technology Challenge Grants), the Division of Innovation and Development in the Office of Special Education Programs, and the Fund for

³ Department of Education staff interviewed included Pat Knight (Assistant Secretary's Office); Ed Fuentes, Naomi Karp, Carol La Campagne, Joseph Conaty, and Deborah Inman (Center Directors); Gilbert Garcia, Jerry Lord, and Beth Fine (At-Risk); Veda Bright (Early Childhood); Duc Le To, James Fox, and Ron Anson (Policy); Delores Monroe (Postsecondary); and Jackie Jenkins, Clara Lawson Holmes, and Judith Anderson (Achievement). Two representatives of professional associations were also interviewed: Jerry Stroufe (AERA) and David Johnson (Federation of Behavioral, Psychological and Cognitive Sciences).

Innovation in Postsecondary Education. Other federal agencies included the National Institutes of Health, the National Science Foundation, and the National Endowment for the Humanities.

Organization of This Report

Chapter 2 of this report reviews the specific requirements embodied in the new standards. Chapter 3 describes the FIS and center competitions conducted in fiscal years 1996 and 1997. Chapter 4 presents the study findings regarding the fit between reviewers and the applications they reviewed, the quality of the reviews performed, and the review process of the six panels examined in depth. Chapter 5 presents conclusions and recommendations of the study. The appendices include background material for the study: Appendix A is a copy of the OERI/NERPPB standards governing the conduct and evaluation of grants and cooperative agreements; Appendix B is the Technical Review Form; and Appendix C provides the evaluation criteria and specific factors used to rate applications for FIS and center competitions.

THE NATURE OF THE STANDARDS

Before examining the implementation of the new OERI/NERPPB standards for conducting application reviews, it is important to understand what those standards require. This chapter describes each important element in the standards and reflects briefly on associated implementation issues that have arisen. The elements described include the qualifications and selection of peer reviewers who review applications for grants and cooperative agreements; the rules for conducting the competitions; and the process for evaluation, ranking, and selection of award recipients.

Qualifications of Peer Reviewers

The standards require that individuals selected as peer reviewers have the following qualifications: "i) Demonstrated expertise, including training and experience, in the subject area of the competition. ii) In-depth knowledge of policy or practice in education. iii) In-depth knowledge of theoretical perspectives or methodological approaches in the subject area of the competition." Because there has been considerable confusion about who can serve as a peer reviewer, these requirements are quoted as they appear in the regulations (*Federal Register*, September 14, 1995:47811). All three conditions would appear to apply to *all* individuals who review applications; there is nothing in the standards to suggest that it is sufficient for the reviewer panels to reflect the three qualifications collectively.

The wording of the first qualification indicates that determining the subject area of a

competition is critical to determining who is qualified to serve as a peer reviewer. A major issue that has arisen in practice is determining whether the "subject area" means "research about the subject area." If the competition is the field-initiated studies (FIS) Educational Research Grant Program, the "subject area" would appear to be research within the purview of the institutes participating in the program (e.g., early childhood education, education of at-risk children). The second qualification, in-depth knowledge of policy or practice, seeks to ensure that, in addition to expertise in the subject area of the competition, all reviewers have some grounding in the practice of education. The importance of research expertise would appear to be reinforced by the third qualification—in-depth knowledge of theoretical perspectives or methodological approaches in the subject area of the competition. This requirement goes further than simply requiring expertise in a field; it specifies two important components of expertise and indicates that all reviewers must have at least one of them. All reviewers must come to the table with "in-depth knowledge" of theory or methods in the area of the competition. One of those components—in-depth knowledge of methods—is a subset of research expertise. The other—in-depth knowledge of theory—might exist apart from research expertise, but the likelihood of an individual having in-depth theoretical understanding without a research background would be small.

There is little doubt that considerable effort is required to find individuals who meet all three qualifications in the standards. Nonetheless, the standards should be, and indeed appear to be, an effort to raise the bar, ensuring a high-quality group of peer reviewers who will be able to provide expert reviews and lend insight and stature to the peer review efforts of OERI. By requiring that all members of a panel meet all three qualifications, the standards clearly

indicate that creating a panel is not to be a "mix-and-match" effort, with one reviewer representing the subject area, one practice, and one methods or theory, as some have described it. Rather, the standards indicate that all peer reviewers should possess all three qualifications.

Selection of Peer Reviewers

In addition to the individual qualifications for peer reviewers, the standards require that OERI (the Secretary) select "to the extent feasible...peer reviewers for each competition who represent a broad range of perspectives." This requirement is not further elaborated in the discussion of the rules, so it is not immediately apparent how it should be implemented. In the context of research applications, a broad range of perspectives could mean reviewers who differ with regard to disciplinary background within a broad area such as policy or at-risk children (e.g., education psychologists, ethnographers, sociologists), theoretical orientation, methodological approach, research role (academic researchers, contract researchers, voluntary sector researchers), or any number of other characteristics. What is clear from this requirement is the desire to avoid a situation in which all reviewers of an application share the same viewpoint on what is legitimate, appropriate, or feasible in a particular research field.

Conduct of the Competitions

The standards also prescribe certain aspects of the competition, including when peer review is to be used, the minimum numbers of reviewers of an application, reviewer obligations, the review process, and evaluation criteria. First, the rules specify that peer review is to be used for the review and evaluation of all applications for grants and cooperative

agreements that exceed \$100,000.¹ Later in the regulations, however, the prescription appears to vary somewhat, with the language noting that fewer than three reviewers may be used for grant and cooperative agreement awards under \$50,000 "if the Secretary determines that adequate peer review can be obtained using fewer reviewers." This clause appears to extend downward the purview of peer review, but it may be that peer review is discretionary at this funding level. Nonetheless, the clause adds the important criterion of at least three reviewers for grant or cooperative agreement awards above \$50,000. In addition, awards above \$1,000,000 require at least five peer reviewers.

The standards also specify a set of broad evaluation criteria and specific factors from which OERI may select specific items appropriate to each grant or cooperative agreement competition. The broad criteria include national significance, the quality of the project design, the quality and potential contributions of personnel, the adequacy of resources, and the quality of the management plan. The specific factors offer a wide range of options; some are oriented to research competitions, while others are more reflective of demonstrations or program grants. The rules allow for complete discretion with respect to which of these broad criteria and specific factors are used in any competition.

The review process is spelled out in some detail. First, reviewers "must be given a number of applications to evaluate," although no specific number is cited. This requirement may have been inserted to give the reviewers a better perspective on the quality of applications than would be provided by a single review. The rules here appear to acknowledge that

¹ Peer review is also required for contracts exceeding \$100,000, but this report does not address contract administration.

ultimate decisions to award grants are based on relative rankings of applications.

Reviewers are instructed to "independently evaluate" and rate (i.e., score) each application. Evaluations and ratings are to be based on the applicable evaluation criteria and weights assigned to each criterion. The evaluations are to be accompanied by "concise written comments based on the reviewer's analysis of the strengths and weaknesses of the application with respect to each of the applicable evaluation criteria." After the independent evaluation/rating, reviewers who evaluated "a common set" of applications are to convene and discuss the strengths and weaknesses of those applications. "Each reviewer may then independently reevaluate and re-rate an application with appropriate changes made to the written comments." After this process has been completed, "reviewers shall independently place each application in one of three categories, either "highly recommended for funding," "recommended for funding," or "not recommended for funding."

The emphasis throughout the discussion of the review process in the standards is on independent judgment by reviewers. First, the evaluation of applications is to take place prior to meetings among reviewers. After the discussion of application strengths and weaknesses, each reviewer is to make an independent reassessment and revise his or her scores and written comments. Finally, after the review of all applications has been completed, reviewers are to independently assign each application to one of the above three categories (i.e., highly recommended, recommended, not recommended). In essence, the discussion process is an opportunity for the reviewers to gain additional information, perhaps a sense of different viewpoints and perspectives, on the proposed research. Given the emphasis on individual decision making, the review is clearly not the point at which collective decisions are made on

application ratings or desirability for funding.

Evaluating, Ranking, and Selection of Award Recipients

Once the peer review process has been completed, several important steps must occur before awards are made. First, "the Secretary prepares a rank order of the applications based solely on the peer reviewers' ratings." Then, "the Secretary determines the order in which applications will be selected for grant and cooperative agreement awards." The standards require the Secretary to consider a wide array of information, including "(1) An applicant's ranking. (2) Recommendations of the peer reviewers with regard to funding or not funding. (3) Information concerning an applicant's performance and use of funds under a previous Federal award. (4) Amount of funds available for the competition. (5) Any other information relevant to a priority or other statutory or regulatory requirement applicable to the selection of applications for new awards." Clearly, the Secretary has considerable latitude with respect to making awards once the peer review process has been completed.

There are several important features of this prescription for making awards. First, peer reviewer rankings are important, but they are by no means the only consideration in making awards. Second, previous performance may be considered, but the rules are not explicit about who makes such an assessment or how. Third, limited funding may be considered, and could presumably play several roles: it could result in passing over a single expensive project, but it could also result in deciding to support more (or fewer) projects overall, shifting awards as a whole to a different set of projects than might have been funded had rankings alone been considered. Finally, statutory priorities or other statutory requirements can be invoked to

support applicants that may not have received awards based solely on rankings.

Procedures Not Specified by the Standards

It is also important to note what is not covered by the standards. Among those items are (1) priorities that may be identified in program announcements or application packages, but are not found in statutes or regulations; (2) modification of reviewer rankings through statistical manipulations in order to standardize the rankings; (3) the efficacy of multistage reviews; and (4) use of standing panels.

Within OERI, priorities that do not carry the weight of law or regulation are often cited in application packages. While the text of the package may point out that the priorities are nonbinding, it sometimes remains unclear how reviewers and others should view those priorities in deciding on application strengths/weaknesses, rankings, and awards.

With respect to standardization of scores, some grant competitions and some institutes have elected to use a Department of Education (ED) standardization process in creating their rankings, while others have not. This process is designed to correct for possible bias introduced by different reviewers' approaches to assigning raw scores (i.e., some reviewers tend to score high, while others tend to score low). The standardization process is based on certain assumptions about the distribution of applications and the behavior of reviewers. According to Department of Education information, the standardization process assumes

that the varying quality of applications in the entire pool of applications is normally distributed, (i.e., a similar number of good, average, and poor

applications; the applications distributed to each panel are normally distributed; all panels have the same training and direction; and any resulting deviation is due to reader bias).

Multistage reviews—which usually entail a first-stage review by peers, followed by a second-stage review by peers and/or a combination of peers and other stakeholders—are commonly used by other federal research agencies, often to separate judgments about research quality from those about policy significance. Standing panels comprising individuals who meet the standards for qualified reviewers and serve specified terms, generally for several consecutive years, are also used by other agencies. While such reviews and panels are not addressed in the standards, the Secretary's discretionary powers in making grant awards (especially with regard to obtaining other information on priorities or requirements) would appear to make second-stage reviews and standing panels possible if they were undertaken to assist the Secretary in decision making, and if the applicants were aware that this process was being employed. Second-stage reviews and standing panels are so widely used by other federal research agencies (such as the National Institutes of Health and the National Science Foundation) that it is certainly worth considering further whether they are allowed by the OERI/NERPPB standards.

DESCRIPTION OF THE COMPETITIONS

Field-Initiated Studies Research Grant Program, Fiscal Years 1996 and 1997

This section describes the procedures used for the field-initiated studies (FIS) competitions in fiscal years (FY) 1996 and 1997. It draws on information presented in the technical review plans and slate recommendations for the competitions, as well as from Department of Education staff interviews.

Background

Under Section 931 of Title IX of the Educational Research, Development, and Improvement Act of 1994 (Public Law 103-227), OERI held a discretionary grant competition for FIS grants. "Standards for the Conduct and Evaluation of Activities Carried Out by the Office of Educational Research and Improvement" (34 Code of Federal Regulations [CFR] Part 700), provides program criteria for applicants' proposals and guidance for the peer review process.

In general, each institute may fund only applications for studies that are within its legislative mission. In the application package, applicants were instructed that they were to apply to one of five institutes. Applicants were also invited to address one of the priorities cited in the Office of Educational Research and Improvement's *Research Priorities Plan*.

However, it was clear from the application package that addressing these priorities was not mandatory. Specifically, the text states: "Applicants that address one of these priorities will not receive competitive or absolute preference over other applicants."

Generally, grant funds support research projects averaging \$150,00 per year and from 1 to 3 years in duration. In FY 1996, 45 awards were made from among a total of 647 applications. In FY 1997, 30 awards were made from among a total of 532 applications.

Compilation of Reviewer Rosters

In accordance with the OERI/NERPPB standards, each institute's work team,¹ in collaboration with the institute directors, prepared two lists of reviewers prior to the start of the panel reviews—a primary list and a list of alternates. If a reviewer from the primary list could not serve, another reviewer from the approved list of alternates was to be invited. The lists of primary and alternate reviewers were submitted to the institute directors. In FY 1997, about 30 percent of those on the primary list had served as reviewers in FY 1996, with the remainder being new reviewers. The list of alternate reviewers included available current FIS project directors.

According to OERI staff, reviewers were selected who had subject area expertise in the areas in which the institutes anticipated receiving proposals, according to their mission statements and the topic areas included in the *Research Priorities Plan*. Some institutes waited until the applications had been received and they knew the application topics before

¹ Effective October 1, 1997, the Department of Education implemented a revised grant-making process. One aspect of the revised process was that grant slates were to be compiled and approved by collaborative work teams (teams composed of individuals from a given institute).

constructing the review panels. Others formed preliminary panels based on anticipated application topics. The institute directors met to ensure that there would be no overlap in reviewers across institutes.

A file containing information about potential reviewers is maintained in the program office; it shows the reviewers' qualifications and availability to participate in the review process. The availability status of reviewers is updated as staff begin to recruit reviewers for each competition. In FY 1996, the Department published a request for reviewers in the *Federal Register* and in the *OERI Bulletin*. The need for qualified reviewers is announced to professional associations at annual meetings, through departmental announcements, and through direct recruitment by program staff.

Screening of Applications

Department of Education staff screened applications for eligibility and conformance with the regulations governing the program, and reviewed them for completeness. Applications were rejected if they did not meet the statutory eligibility requirements under the authorizing statute or Education Department General Administrative Regulations (EDGAR), if they were not sufficiently complete, if the applicant was not eligible, or if there were no funds available within OERI to address the subject of the application. In FY 1996, 72 applications were deemed ineligible; in FY 1997 the number of ineligible applications was 56.

Conflicts of Interest, Assurances, and Waivers

Reviewers were required to sign one of two forms, depending on whether they were to

be paid for their services.² Reviewers who reviewed and evaluated applications at no cost to the government were required to sign a Gratuitous Services Agreement form. Any reviewers who were to be paid travel expenses, per diem, and an honorarium were required to sign an Agreement for Grant Application Reviewers Who Receive Compensation form, as well as a Drug-Free Workplace Certification form.

These forms also contain a "Conflict of Interest" section. Conflict of interest is defined in 34 CFR Part 700, "Standards for the Conduct and Evaluation of Activities Carried Out by Office of Educational Research and Improvement," Section 700.13. Each institute was to submit a request for blanket waivers of conflict of interest when a significant majority of reviewers were affiliated with institutions that had submitted applications, but were not directly associated with those applications. Individual blanket waivers were provided in FY 1996. In FY 1997, to cover cases in which reviewers were employees of a university but did not have direct involvement with applications from that university, a joint request (covering all five institutes) for a blanket waiver of conflict of interest was signed by Ramon Cortines, Acting Assistant Secretary, on June 16, 1997.

At the beginning of the review process in FY 1996, a list of assigned applications was sent to each reviewer to allow identification of any conflict of interest as defined by instructions. Department of Education staff contacted reviewers to confirm their receipt of this list and to ensure that there were in fact no conflicts of interest. In FY 1997, an outside contractor sent reviewers the applications; reviewers were told to contact the contractor to confirm receipt of the applications, and to contact their panel's chair immediately if there were

² Federal employees are not paid for serving as peer reviewers. In some cases, funds are unavailable for reviewers regardless of their affiliation.

a conflict of interest or other problem. (In both years, during recruitment, staff also asked reviewers to inform them about any potential conflicts of interest.) Reviewers were also asked to sign the "Conflict of Interest" section of the relevant form, stating that no conflict existed and that if one arose, they were obligated to and would inform the program official. If a reviewer had a conflict of interest involving any application in the competition, that reviewer was not to read at all unless a waiver had been signed by the Office of the General Counsel.

Orientation Instructions for Reviewers

FY 1996. An orientation package was sent to reviewers. It included a description of the program purpose and key review dates, a copy of the application package, the Technical Review Form (both hard copy and diskette), and a copy of the Horace Mann Learning Center workbook for reviewers. A second package was subsequently sent out, containing instructions for reviewers; a list applications assigned to panel members; and a Gratuitous Services Agreement form, or an Agreement for Grant Application Reviewers Who Receive Compensation and Drug-Free Workplace Certification form.

FY 1997. An independent contractor was hired to assist with the logistics for the competition. The contractor sent a preliminary letter to reviewers to confirm the dates of the panel meetings and explain hotel and travel arrangements to out-of-towners and locals. Approximately 3 weeks before the review panels were convened, each confirmed reviewer received a package containing a thank-you letter; reviewer instructions (including travel and hotel specifics); a Technical Review Form and disk (same for all institutes, but with institute-specific headings); a copy of the Horace Mann Learning Center workbook for reviewers; a list

of assigned applications and panel members; assigned applications; and a Gratuitous Services Agreement form, or an Agreement for Grant Application Reviewers Who Receive Compensation and Drug-Free Workplace Certification form.

Basis for Assigning Applications to Reviewers

Every application was to be read by a panel of three reviewers. Applications were to be assigned to panels according to the fit between panel expertise and project topics. According to Department of Education staff, to the extent possible, the three reviewers on the panel were also categorized as a specialist in the subject area, a generalist, and a methodologist. To avoid conflict of interest, reviewers were not given applications from institutions or organizations with which they were affiliated or with which they had any financial or working relationship (past or present). Across the institutes, there were 218 reviewers in FY 1996 and 147 in FY 1997.

On-Site Orientation (FY 1997 only)

On the morning of the first day of panel reviews in FY 1997, Department of Education staff provided an orientation for reviewers. Staff from various institutes discussed the review process and showed a film on the peer review process.

Review Process

Reviewers had access only to the applications assigned to their panels, and were not given any information about applications reviewed by other panels. In most cases, panels were

expected to read at least 10 applications. In FY 1996, reviewers rated applications off site and conferenced via telephone only if there were significant discrepancies among the reviewers' scores. In FY 1997, reviewers were asked to complete their preliminary review of applications by mail and then met in Washington, D.C., to complete the evaluations as a panel. Reviewers met as panels of three, and each panel was assigned a chair who was responsible for facilitating the timely discussion of applications. A Department of Education staff member was assigned to each panel to facilitate the panel discussion and answer reviewers' questions. Reviewers discussed their reviews, generally one application at a time, and following the discussion were allowed to adjust their scores and revise their comments accordingly. They were instructed to make score adjustments independently of other reviewers. At the conclusion of the review process, each panelist was to place each application in one of three categories: highly recommended for funding, recommended for funding, or not recommended for funding.

Ranking of Applications

Following completion of the panel review, the program office entered the scores of each reviewer for each application on a panel summary form. Average raw scores were then computed from the reviewers' individual raw scores. These raw scores were entered into each institute's database to establish an initial rank order. In cases in which application scores were standardized (see below), the department's Application Control Center used these raw scores for that purpose.

Standardizing of Scores

In FY 1996, the Policy Institute used the Department of Education discretionary grant application score standardization process; all other institutes used raw scores. In making 1997 FIS grant awards, four of the five institutes used the score standardization process; the exception was the Institute on Educational Governance, Finance, Policymaking, and Management.

Making of Grant Awards

According to Department of Education staff, their funding recommendations were consistent with the results of the panels' technical reviews, and the highest-rated applications were selected for funding, with one exception. In FY 1996, staff deviated from the panelists' rank ordering because the work proposed in one application duplicated work carried out by the applicant with other OERI funds. In two other instances, staff selected one of two equally rated applications (both could not be funded given available resources) based on their judgment as to which would contribute most to the field.

Funding slates were prepared that contained the funding recommendations of the collaborative work team, the institute directors, and representatives of the Office of the Assistant Secretary (OAS), along with the rank-ordered list of all applications reviewed. The institute directors and team members signed the slate of recommendations. After the awards had been announced, unsuccessful applicants were sent regret letters signed by the work team leader, along with copies of the review forms.

Program office staff were to send original unfunded applications and related documents

to the Federal Records Center within 90 days after awards had been made and unsuccessful applicants had been notified. These files are to be retained by the Federal Records Center for 3 years. Any other copies of the applications and related documents are to be destroyed in a manner that preserves their confidentiality.

National Educational Research and Development Center Program Competitions

Overview

The Educational Research, Development, and Improvement Act of 1994 required OERI to continue to fund centers, using at least one-third of the funds available to each institute in any fiscal year for this purpose. Prior to the FY 1996 competition, the first to be conducted under the new OERI/NERPPB standards, Department of Education staff met to resolve three key issues: the number of centers to fund, funding levels for each, and topic areas to be addressed by the new centers.

Department of Education staff developed priorities for seven new research centers. They did so through a lengthy process that involved notices published in the *Federal Register* for public comment, OERI staff work group papers, meetings with national educational associations and organizations, public meetings in five states, expert panel reviews, NERPPB discussions, and internal Department of Education staff reviews. The seven priorities were as follows: Priority 1, Enhancing Young Children's Development and Learning; Priority 2, Improving Student Learning and Achievement; Priority 3, Improving Student Assessment and

Educational Accountability; Priority 4, Meeting the Educational Needs of a Diverse Student Population; Priority 6, Improving Postsecondary Education; and Priority 7, Improving Adult Learning and Literacy. Each center was to conduct a coherent, sustained program of research and development, and address problems and issues of national significance in its individual priority area.

In FY 1996, seven distinct but coordinated competitions were held to fund centers in the above seven priority areas.³ The program competition was announced in the *Federal Register* on September 14, 1995, and the closing date for receipt of applications was December 15, 1996. Government furloughs, fear of congressional rescission of center funding, and a snowstorm in January resulted in later closing dates for receipt of applications and caused delays in the review process for several of the centers.

In FY 1997, two distinct competitions were carried out—Research on Policy and Teaching Excellence and Research To Improve Children's Early Reading. The first was announced in the *Federal Register* on April 11, 1997, with a closing date of June 6, 1997; the second was announced on March 26, 1997, with a closing date of May 28, 1997.

In FY 1996, 8 center awards were made from a total of 38 applications. In FY 97, 2 center awards were made from a total of 9 applications.

Composition of Review Panels

In FY 1996, OERI staff established review panels consisting of approximately 5-10 peer reviewers for each of the competitions, depending on the expertise needed. Peer

³ Under 34 CFR Parts 74, 75, 77, 80, 81, 82, 85, 86, 98, 99, and 700.

reviewers were to possess the qualifications specified by the standards (see Chapter 2). Names of potential reviewers were generated by database searches and staff recommendations; Department of Education staff contacted grantees and examined past lists of reviewers. Nonfederal reviewers were paid an honorarium of \$750.

In FY 1997, the Policy Institute prepared a primary list of reviewers and a list of alternates, selected from a comprehensive file of reviewers maintained in the program office and from work team recommendations. For the competition to award the Center on Early Reading, five potential reviewers were identified by the Achievement Institute's work team and two by Early Childhood Institute staff.

Conflicts of Interest, Assurances, and Waivers

As with the FIS panels, reviewers were required to sign one of two forms, depending on whether they were to be paid for their services.

Reviewers were informed early and often about the requirements regarding conflict of interest. They were asked to notify OERI staff immediately if their institution was planning to apply for a center award for which they had been asked to serve as a reviewer, or if they or a family member would benefit from the outcome of the competition. To assist in identifying any conflicts of interest, each reviewer was asked to examine the list of applications assigned to his or her panel prior to the beginning of the review, and OERI staff reviewed applications before mailing them to reviewers.

Screening of Applications

As with FIS applications, Department of Education staff screened applications for eligibility and conformance with the regulations governing the program, and reviewed applications for completeness.

Technical Review Form

The application Technical Review form was used to review applications (see Appendix B).

Orientation for the Panels

Prior to each panel meeting, peer reviewers received by mail a general letter of instructions, a set of reviewing instructions, a copy of the application package, copies of the assigned applications, and the application Technical Review Forms (both hardcopy and electronic versions). The letter for reviewers was uniform, except for references to the specific priority area for each panel. The letters and materials were sent out under the signature of the acting director of the institute responsible for the panel. Institute staff contacted peer reviewers by telephone and arranged for one reviewer on each of the panels to serve as chair. In some cases institute staff assigned responsibility for presenting and discussing given applications to specific reviewers on a panel.

On the first morning of the panel meetings, there was an orientation session that included a representative from the Grants and Contracts Office. During this meeting, OERI staff briefly reviewed the selection criteria and stressed the need for full documentation of all scores. Reviewers were reminded to evaluate each application independently of each other,

evaluate the application based on an assessment of its quality according to the evaluation criteria and the weights assigned to those criteria, and support the rating assigned to the application with concise written comments based on an analysis of the strengths and weaknesses of the application with respect to each of the applicable evaluation criteria.

Conduct of the Review Panel

A one-tier review was conducted in three phases. First, all reviewers reviewed and scored all applications received for the priority area they were addressing according to the established selection criteria. Second, reviewers on each panel met as a group in Washington, D.C., and discussed applications assigned to their panel. At the beginning of the meeting, preliminary scores were posted for all applications. All of the applications were discussed, and discussion was concluded on those applications whose preliminary scores, ratings, and evaluations were very weak. The panelists again reviewed the remaining applications. Next, each reviewer had the opportunity to independently reevaluate and re-rate any of the applications and make changes to the written comments as appropriate. The applications were then discussed to share information, not to seek consensus. Third, the rank order of the applications—based solely on the reviewers' scores—was calculated and posted. In some cases, reviewers identified critical issues for the highest-ranking applicants. These issues were used by OERI staff to formulate questions for clarification that were transmitted to applicants for response before awards were made.

Department of Education Staff Role

Draft Summary 1/30/99

OERI staff provided an orientation for reviewers prior to the review, attended panel meetings, answered any questions that arose, and examined review forms for accuracy and completeness. They also prepared clarification/negotiation questions for potential awardees. In addition, they prepared funding recommendation memoranda for the Assistant Secretary's approval.

Selection of Applications for Award

The Assistant Secretary selected applications for award, considering the applicants' ranking by the peer reviewers, recommendations of the peer reviewers with regard to funding or no funding, information concerning applicants' performance and use of funds under a previous federal award, the amount of funds available for a competition, and any other information relevant to a priority area or regulatory requirement applicable to the selection of applications for new center awards. In no case did the Assistant Secretary's selection deviate from a panel's rank ordering.

Applications Not Selected for Funding

Applicants whose proposals were not selected for funding received a letter following the congressional notification of funded applicants and the award of the cooperative agreements.

Specifics of the Competitions, FY 1996

Priority #1. The National Institute on Early Childhood Development and Education

supports the Research and Development Center on Enhancing Young Children's Development and Learning, University of North Carolina, at \$2.75 million per year for 5 years. For this competition, seven applications were submitted for consideration by a panel of five reviewers.

Priority #2. The National Institute on Student Achievement, Curriculum, and Assessment supports the National Research and Development Center on Improving Student Learning and Achievement in English, the University of Albany, State University of New York, at \$2.5 million a year for 5 years. The National Institute on Student Achievement, Curriculum, and Assessment also supports the National Research and Development Center on Achievement in School Mathematics and Science at the Wisconsin Center for Educational Research, University of Wisconsin, for \$2.5 million a year for 5 years. For this competition, six applications were submitted, ten reviewers participated, and two awards (referred to above) were made.

Priority #3. The National Institute on Student Achievement, Curriculum, and Assessment supports the National Research and Development Center for Student Assessment and Educational Accountability, the University of California, Los Angeles, at \$2.8 million per year for 5 years. For this competition, two applications were reviewed by a panel of seven reviewers.

Priority #4. The National Institute on the Education of At-Risk Students supports the National Center for Research on Education, Diversity, and Excellence, the University of California, Santa Cruz, at \$4 million per year for 5 years. For this competition, eight applications were submitted for review, and five external reviewers participated.

Priority #5. The National Institute on Educational Governance, Finance,

Policymaking, and Management supports the Consortium for Policy Research in Education, the University of Pennsylvania, at \$2.8 million per year for 5 years. For this competition, five organizations submitted applications that addressed the priority "Increasing the Effectiveness of State and Local Education Reform Efforts," and six reviewers evaluated applications.

Priority #6. The National Institute on Postsecondary Education, Libraries, and Lifelong Learning supports the National Center for Postsecondary Improvement, Stanford University, at \$2.5 million per year for 5 years. For this competition, six proposals were submitted and were reviewed by a panel of five individuals.

Priority #7. The National Institute on Postsecondary Education, Libraries, and Lifelong Learning supports the National Center for the Study of Adult Learning and Literacy, Harvard University/World Education, at \$2.5 million per year for 5 years. For this competition, eight proposals were submitted, and a panel of nine reviewers met to discuss the applications.

Specifics of the Competitions, FY 1997

National Center on Early Reading. The Student Achievement Institute and the Early Childhood Institute support the Center for the Improvement of Early Reading, the University of Michigan, at \$2.3 million per year for 5 years. Five applications were submitted for review, and seven external reviewers served on the panel.

National Research Center on Policy and Teaching Excellence. The Policy Institute supports the National Center on Policy and Teaching Excellence, the University of Washington (in collaboration with four other universities), at \$1.5 million for 5 years. Four applications

were submitted for review, and seven external reviewers served on the panel.

FINDINGS

This chapter summarizes study findings in three areas: the substantive fit between reviewers and applications; the quality of the peer reviews; and a more indepth examination of the review process of six panels, including four fiscal year (FY) 1997 field-initiated studies (FIS) competition panels nominated by institute staff as particularly successful and two center competition panels—one nominated as successful and one as problematic.

Substantive Fit Between Reviewers and Applications

Methods

To determine how well the peer reviewers reflected the intent of the standards, we conducted an exercise to match application content, theory, and methods with the background and experience of the individuals who conducted the reviews. Because of time and data collection constraints, data on peer reviewers' credentials were limited to the resumes they submitted to OERI at the time of the reviews. While it would have been desirable to obtain additional and more detailed background information, in many, if not most instances, those resumes were also the main data available to the OERI staff in making reviewer selections. In a few cases, the documentation submitted by the reviewer was not a resume (e.g., a press release or bio) or was an abbreviated resume.

For each of the panels and applications selected through the stratified random selection procedure described in Chapter 1, we requested from OERI the resumes of all three peer

reviewers. We then examined the applications and peer reviewer credentials (those we obtained) for 12 FIS panels for FY 1997 (29 applications and 35 reviewers) and 6 FIS panels for FY 1996.¹ We also examined the fit between applications and reviewer credentials for those five center competitions for which we had the necessary information. We then constructed data displays resulting from this inquiry, in which the content of each application and data on peer reviewers' education, current position, and research experience were briefly summarized. For each application a short statement about the fit between the two was prepared.

In general, we used the following approach in our assessment. If an individual had a doctorate, we looked at the field of the doctorate and the individual's publications. If the field generally required research for a doctorate (e.g., education psychology or education research), we assumed that the individual had a research background. If the field might or might not require original research (e.g., curriculum and instruction, educational administration), we looked at positions held and publications as well. We did the same for individuals without doctorates. Sometimes it was impossible to determine whether an individual had a research background because publications were incomplete or titles did not yield sufficient information, in which case that limitation was noted in the assessment. We also found that it was largely impossible to determine expertise in methods or theory from the resumes. The most we could learn was whether the reviewer had conducted roughly similar types of research in the broad subject area of the competition. We did not attempt a detailed match between reviewer credentials and specific applications. Not only was the data insufficient to permit that type of

¹We did not examine one postsecondary FIS panel because one of the authors of this report was a consultant on an application reviewed by that panel.

analysis, but we recognized that to do so would impose a much stricter rule than is suggested in the standards.²

Reviewers were also interviewed regarding possible concerns about serving as reviewers (e.g., lack of knowledge of the subject area and/or methods, conflict of interest, timing of reviews); whether the concerns had been addressed satisfactorily by Department of Education staff; the extent to which the subject area of the competition was described in sufficient detail for them to determine whether they were qualified to review applications; and their assessment of their own and their fellow panel members' qualifications for serving as reviewers (e.g., familiarity with subject area, proposed methods, scope of the design, etc.).

Findings from the Review of Resumes

While most of the reviewers in the sample had conducted research in education, a sizeable minority had not. We focus here on the FY 1997 FIS reviewers, the group for which the greatest amount of systematic information was available. Of the 35 reviewers on 12 panels whose resumes were reviewed, 17 appeared to be educational researchers, and an additional 6 may well have had research experience, although the resumes for these individuals were insufficient to make that determination (e.g., their resumes showed they had doctorates, but their publications were missing). The remaining 12 individuals (about a third) did not indicate any research experience or publications on their resumes. They included persons who had served as teachers, school administrators, state officials, tribal officials, teacher trainers, and

² An OERI staff member has pointed out that reviewer credentials may not match applications when the applications are far afield from the institute's focus. This is undoubtedly the case in some instances, but because we were not seeking a perfect fit between reviewers and applications, this problem was not a major issue for our review.

university administrators. Most had a solid background in education policy or practice and familiarity with the general subjects being reviewed (early childhood education or science education), but did not meet the criterion of having studied and conducted research in the general field in which they were reviewing applications.

Among the 23 individuals who had or may have had research training and experience, most had that experience in broad areas related to the competitions. For example, individuals with backgrounds in early childhood education were likely to review applications in that area. Furthermore, subject area fit often extended to a more detailed level, for example, with individuals knowledgeable about science education reviewing science education applications within a larger field (student achievement). This level of fit was not always the case with respect to methods, however. Thus, for example, individuals who had little or no experience in studying large-scale program or policy implementation evaluated applications aimed at studying the longitudinal effects of curriculum or policy reforms.

Across institutes, perhaps the most common area in which research experience appeared to be lacking was the design and conduct of evaluations. Of the 29 applications we matched with reviewers, 10 were evaluation studies in whole or in part. These applications proposed studies ranging from small-scale experimental design tests of new curricula to large-scale testing of interventions with nationally representative samples of children, teachers, or others. Judging from their experience and publications, few of the individuals who reviewed those applications appeared to have conducted evaluations themselves, let alone experiments or studies requiring elaborate sampling designs. The reviewers selected were likely to be familiar with the subject of the evaluation (early childhood education or at-risk youth), but they did not

appear to be as familiar with the types of studies commonly proposed--studies using experimental or quasi-experimental designs to conduct evaluations.

Saying that up to 23 of 35 reviewers had training and/or experience in the research areas of the competition does not imply that they were all experts in those fields, however. Of the 17 who were clearly education researchers, two indicated that their Ph.D.s were (or would be) earned in 1997, meaning they may not have had those degrees at the time they reviewed the applications. In addition, several of the reviewers with research experience had only limited amounts of experience in research; we counted as researchers individuals with a research dissertation or a few research publications. The standards speak of "expertise," but it is impossible to make that assessment from resumes.

Most of the panels for the center competitions for which resumes were available appeared to be qualified to review the center applications. Not all the members of each panel held doctorates or had conducted research in the competition subject area, but most on each panel met these criteria. The exceptions among the five panels were (1) the Postsecondary Improvement Center, for which most of the reviewers had conducted some research, but only a minority appeared to have studied reform in postsecondary education (although two had extensive backgrounds in research on occupational training); and (2) the Adult Literacy Center, for which only a minority of the reviewers had either doctorates or research experience. We were provided with six of the nine reviewers' resumes, so it is possible that the data on the missing reviewers would alter this conclusion.

Findings from the Interviews

Draft Summary 1/30/99

Of the 14 reviewers surveyed who had participated in the FY 1996 FIS competition, 12 had no concerns about serving as a reviewer at any time in the process (question 1). One had concerns initially about serving because of a potential conflict of interest, but those concerns were adequately addressed by Department of Education staff. One reviewer was concerned about his/her lack of knowledge in the specific subject area of the competition and about the timing of the competition (not having adequate time to prepare for the review), but was told by Department of Education staff that there was no problem with regard to subject matter knowledge because he/she had been recruited for policy rather than specific subject matter expertise.

All but 1 of the 14 reviewers said the subject area was described in sufficient detail for them to make a determination as to whether they were qualified to serve (question 2).

With regard to reviewers' assessment of their own expertise (question 3), 13 reviewers reported that their expertise was appropriate. One expressed concern because of lack of specific expertise in the subject areas of the range of proposals to be reviewed.

Reviewers' assessment of the expertise of their fellow panelists was mixed (question 4). Of the five panelists who participated in panel discussions (FY 1996 was mostly a mail-out review, with panel discussion only for applications with discrepant ratings), two reviewers found their fellow panelists' expertise satisfactory, while three were more critical of the expertise of their fellow reviewers. Of these three, two cited lack of expertise in research design and methodology as their concern. One reviewer stated that a fellow panelist lacked objectivity, had his/her own "agenda", and had not devoted proper time to reviewing applications prior to the panel meeting. One reviewer was concerned because the other

reviewers lacked practical experience and were not knowledgeable about the subject area of the applications.

Of the 26 reviewers surveyed who had participated in the FY 1997 FIS competition, 22 had no concerns about serving as a reviewer at any time in the process (question 1). Three had concerns initially about serving because of a potential conflict of interest, but their concerns were adequately addressed by Department of Education staff. One reviewer was concerned about his/her lack of research knowledge but was told by Department of Education staff that this was not a problem because at least one other member of the panel would have this expertise.

Almost all reviewers (23) stated that the subject area of the competition was described in sufficient detail for them to make this determination (question 2). One reported there was sufficient information once the materials had arrived. One reported there appeared to be enough information initially, but in retrospect (after completing the reviews), he/she realized there was not sufficient information. One stated there was not sufficient detail.

With regard to reviewers' assessment of their own expertise (question 3), 16 reviewers reported their expertise was appropriate. Two reported it was appropriate because they were working on a panel with other reviewers who complemented their expertise. Three expressed concerns because of a lack of background in research design and methodology. Three stated that they lacked specific expertise in the subject areas of the range of proposals they had to review. Two reported that their expertise was for the most part appropriate.

As in FY 1996, reviewers' assessment of the expertise of their fellow panelists was mixed (question 4). Seventeen reviewers expressed satisfaction. In fact, they described the

panel expertise in glowing terms--"a tapestry of expertise that overall covered the area very well," "the team complemented each other," "each brought different strengths," "outstanding team expertise," and "other reviewers had research background that complemented my practical experience." These comments indicate that reviewers made their assessment based on expertise across the three reviewers, rather than on each reviewer's having expertise in all areas (as specified by the standards; see Chapter 2). Nine reviewers were more critical of the expertise of their fellow reviewers. Seven of these nine cited a lack of expertise in research design and methodology as their concern. One reviewer was concerned because a fellow panel member had not read the proposals prior to the review, but rated them anyway, and one reviewer provided no reason for his/her assessment of a lack of expertise among other panelists.

Of the 14 reviewers who participated in the center competitions, 9 had no concerns about serving as a reviewer at any time in the process (question 1). Five had concerns because of the short lead time they had been given prior to the review.

All but 1 of the 14 reviewers said the subject area was described in sufficient detail for them to make this determination (question 2). One reported there was insufficient information but prior knowledge of the center activities helped. With regard to reviewer assessment of their expertise (question 3), all reviewers reported that their expertise was appropriate.

Reviewers' assessment of the expertise of their fellow panelists on the whole was quite positive (question 4). Eleven reviewers were very satisfied with this expertise, and many commented positively on the "variety of points of view, backgrounds, or perspectives of the panel." Three panelists expressed concern about their fellow panelists because of their lack of

knowledge of research design and methodology, subject area knowledge, bias, not having read the applications prior to the meeting, or modest participation during the panel meeting.

Evaluation of Reviews

Methods

To determine their quality, we read 141 reviews produced for sampled applications in the 1996 and 1997 FIS competitions, as well as five center competitions. We read at least 2, and usually all 3, reviews for each of the sampled FIS applications.³

To guide our reading and assessment of the reviews, we produced a one-page evaluation sheet listing a series of questions about the review:

- Are comments concise (i.e., does the reviewer provide brief, precise, specific, and persuasive arguments about the design, methods, etc. of the proposed research)?
- Does the review make clear whether the research is or is not likely to yield valid and useful information?
- Are the comments related to the evaluation criteria?
- Are the comments consistent with scores (i.e., high scores are accompanied by largely positive comments, low scores are accompanied by largely negative comments, and mixed comments are provided for mixed scores)?
- Are the comments sufficiently elaborated (i.e., is the reviewer's judgment amply and expertly justified)?
- Are there any additional comments?

Concise comments, comments related to evaluation criteria, and comments consistent with

³ Our sampling plan called for sampling reviewer comments on each application, but institutes often supplied us with all three reviews for an FIS application, so we read all three. For the five center competitions, we read all reviews for sampled applications in one of the competitions, but only the reviews of sampled reviewers in the other four.

scores are all required by the standards. We added the additional questions to provide a more complete picture of the quality of the reviews.

It is important to note the specific evaluation factors selected for the grant competitions (see also Chapter 2). The broad criteria were the same for all competitions and included national significance, the quality of the project design, the quality and potential contribution of personnel, the adequacy of resources, and the quality of the management plan. The specific factors under each criterion were different; they appear in Appendix C.

In all competitions, reviewers were asked to provide written comments under each broad criterion, as well as scores. Comments for each criterion were to be divided into "strengths" and "weaknesses." In the hardcopy version of the review form, one page was provided for each broad criterion, with equal space provided for strengths and weaknesses. In addition, a final page, which appeared to be optional,⁴ gave reviewers an opportunity to make overall comments in support of their recommendations, describing strengths and weaknesses, as well as providing suggestions for improving the project in future submissions.

In addition, applicants were interviewed to determine their assessment of the quality of the reviews they received. They were asked how extensive and useful they found the written comments based on the reviewers' analysis of the strengths and weaknesses of the application with respect to each of the application criteria, and the extent to which reviews demonstrated expertise and familiarity with policy and practice in the field of education, as well as in-depth knowledge of theoretical perspectives or methodological approaches relevant to the subject of the competition.

⁴ We assume it was optional because many reviewers did not fill it out.

Findings from the Review of Applications

In terms of breadth of coverage, most reviews met the letter of the standards in that they provided short comments on each broad evaluation criterion, the comments were related to the evaluation criteria, and the scores reflected the comments. To the extent that concise comments were provided, they were most often found under the first two criteria—national significance and design. The other evaluation criteria received much less comment.

Having noted the breadth of the reviews, it is important to note as well that most provided little depth. With respect to both national significance and project design, most reviewer descriptions of application strengths did little more than identify or document what was included in the application, sometimes accompanied by a summary judgment on its quality, sometimes not. In describing application weaknesses, reviewers were more likely to express independent judgment in the area of design than in the area of national significance. For example, 30 of the FIS reviews included a relatively detailed discussion of design weaknesses; most of these comments drew on independent or personal knowledge, not solely on the application. Although there was less detailed description of weaknesses than strengths under national significance, some reviewers considered such weaknesses in detail. Finally, on the whole, other sections of the reviews (staffing, budget, management plan) generated little detailed discussion.

We also categorized reviews as "good," "bad," or "indifferent" based on their breadth and depth of coverage. Of the 79 reviews conducted in FY 1997, about one-third were good: they were detailed assessments of an application's strengths and weaknesses that displayed the

reviewer's knowledge as it was brought to bear on the application. Another 20 percent of the reviews could be characterized as poor: they misstated or simply did not reflect an understanding of the application content; were so poorly written that it was impossible to know what positions they were taking or advice they were providing; or ignored the research components of research projects, focusing instead on the application's intervention as a program or demonstration. Indifferent reviews accounted for about half of the reviews we read: they were fragmentary, and they listed elements in the design or in the needs or theory sections of applications as strengths, with little support for that designation--sometimes adding a short judgment, sometimes not. They offered little if any independent insight from the reviewer. These reviews may have stated one or two critical points (i.e., weaknesses) for a given criterion, but those points were often minor or provided information any observant reader could offer about omitted or incomplete items, such as "instruments not included," or "objectives not clearly stated." Indifferent reviews rarely provided guidance to the applicants about how to improve their applications.

Although far fewer applications were reviewed for the FY 1996 FIS, the quality of those reviews appeared to be somewhat better than that of the FY 1997 FIS reviews. Of the 21 reviews we examined, almost half could be considered good by our criteria; the remainder were largely indifferent.

While some center competition reviews were detailed, far too many fell into the indifferent category. About half of the 41 center reviews we examined were relatively detailed; the other half were brief, and provided mainly the same types of short descriptive or normative statements as those described above for the FIS reviews. As with the FIS reviews,

the attention to criteria other than national significance and design was minimal. Few reviewers wrote more than perfunctory comments on personnel, resources, and management. Lack of attention to management is of particular concern because centers are often composed of researchers drawn from many institutions and require rigorous management.

Findings from the Interviews

There was a great deal of variability both within panels, across panels in a given institute, and across institutes in applicants' responses to questions posed to them. Nor was there any clear pattern by panel or proposal status--acceptance versus rejection. For this reason, the data is presented by competition only.

Overall, applicant assessment of the reviews was mixed.⁵ In terms of applicants' assessment of the usefulness and extensiveness of the reviews, 8 applicants (of 34) in the FIS competitions gave the reviews low ratings on this criterion, 16 applicants (of 34) gave reviews mixed ratings, and about 10 (of 34) rated them high. Ratings on the same criterion for center applicants were 9 (of 17) poor, 5 (of 17) mixed, and 3 (of 17) high.

The reasons for the FIS and center applicants' negative or mixed assessments were varied and included disagreement with the comments, comments that were considered superficial or irrelevant, no comments about design, lack of examples, comments that were illegible, limited explanation for comments, proposal not carefully read, large discrepancies among reviewer comments, reviewer comments too similar to each other, and summary

⁵ It should be noted that in some cases, applicants did not remember the reviews well enough to comment or could not tell from reviewers' comments whether the reviewer demonstrated expertise.

statements that did not mesh with comments in individual categories (e.g., technical quality, national significance).

In terms of the extent to which applicants considered that reviewers demonstrated appropriate expertise (response to question 2), 21 (of 34) of all FIS applicants rated reviewers low or mixed on expertise, and 9 (of 34) rated them as high.⁶ Among center applicants, 14 (of 17) rated reviewers low on expertise and 3 (of 17) as high.

Applicants gave reviewers mixed or poor ratings for a variety of reasons, including failure to understand the significance or value of the proposal; disagreement with reviewer comments about personnel, time, budget, or national significance; comments that reflected poor understanding of research and methodology; comments that did not address information presented in applications; lack of understanding of the substantive focus of applications; lack of understanding of budget, personnel, and management issues, resulting in comments that were inappropriate; and reviewer having his/her own agenda focused on a particular population. The most prevalent concerns, expressed by nine applicants, related to reviewers' lack of understanding of research design and methodology.

Case Studies

Methods

In addition to our overall picture of the competitions, we examined the review process

⁶ Four applicants did not comment.

of six panels in greater depth. The focus of this in-depth examination was four FY 1997 FIS competition panels nominated by institute staff as particularly successful and two center competition panels--one nominated as successful and one as problematic.⁷ This section of the report summarizes the review process in these six cases from the perspectives of reviewers and applicants. In a sense, this discussion highlights what can be learned from panels that OERI staff consider indicative of good practice in the review of applications (the four FIS panels and one center competition were designated as exemplary by OERI staff), as well as problems that have arisen in carrying out the review of grant applications (one center competition was designated as problematic by staff).

Findings

In general, reviewers found the instructional information they received pertaining to the review useful, but most indicated that the review of applications took far longer than the estimates provided by OERI staff. It was not uncommon for reviewers to spend several days or a week or more reviewing 10 FIS or center applications. Most reviewers said they felt they were qualified to review the applications, although a few nonresearchers were unsure in this regard.

Most reviewers found the panel meetings useful and collaborative. Although several FY 1997 FIS reviewers said the orientation meeting and logistics were confusing, almost all found their individual panel discussions enlightening. As discussed earlier, however, a minority were concerned that other panel members did not have the requisite skills and that

⁷ We did not ask staff to identify problematic FIS panels, and we did not examine one of the successful FIS panels because a member of the review team was an applicant to that panel.

nonresearchers were too impressionable. OERI staff were described favorably by most reviewers.

Applicants were generally positive about the application packages and the evaluation criteria. Some indicated that the invitational priorities played a role in their decision to apply, but others said they did not consider those priorities. Some took issue with page restrictions or other constraints that did not have to do with content.

Applicants were very mixed with regard to the quality of the review comments. Not only were unsuccessful applicants less than positive, but some of the successful applicants did not find the review comments useful. Among the issues noted were very short or cryptic comments, comments that provided inaccurate information, or comments that did not appear to fit with scores or rankings.

STUDY CONCLUSIONS AND RECOMMENDATIONS

As noted in Chapter 1, the charge of this study was to examine (1) whether the OERI peer review standards are appropriate and useful, (2) whether they contribute to fair and high-quality competitions, and (3) how the competitions conducted under the standards have operated and how they might be improved. This chapter begins by examining the areas in which the standards are appropriate and useful and as such contribute to fair and high-quality competitions. It then reports findings on how the competitions have operated under the standards and provides recommendations in eight key areas for improving both the competitions and the standards in eight key areas.

Appropriateness and Usefulness of the Standards

The important elements of the OERI standards (*Federal Register*, September 14, 1995:47808-47813) include the qualifications of peer reviewers who review applications for grants and cooperative agreements, the rules for conducting the competitions, and the process for ranking and selecting applications (see Chapter 2).

With regard to the qualifications of peer reviewers, the panel finds that the standards specifying the individual qualifications of peer reviewers are appropriate and useful because those standards ensure that all reviewers have both research and policy or practice experience. The qualifications specified include "i) demonstrated expertise, including training and experience, in the subject area of the competition; ii) in-depth knowledge of policy or practice

in education; iii) in-depth knowledge of theoretical perspectives or methodological approaches in the subject area of the competition."

With regard to the rules for conducting competitions, the panel finds many of the current provisions of the standards appropriate and useful because they standardize procedures across the institutes and make the procedures explicit. These provisions include the following: rules that specify when peer review is to be used (for the review and evaluation of all applications for grants and cooperative agreements that exceed \$100,000); the criterion of at least three reviewers for grant or cooperative agreement awards above \$50,000; the requirement that reviewers "must be given a number of applications to evaluate"; instructions to "independently evaluate" and rate (i.e., score) each application, and to base evaluations and ratings on evaluation criteria and weights assigned to each criterion; the provision that reviewers must accompany their evaluations with "concise written comments based on the reviewer's analysis of the strengths and weaknesses of the application with respect to each of the applicable evaluation criteria"; and the requirement that after the independent evaluation/rating, reviewers who evaluated "a common set" of applications are to convene and discuss the strengths and weaknesses of those applications, and "each reviewer may then independently reevaluate and re-rate an application with appropriate changes made to the written comments."

With regard to the steps that must be taken before awards are made, the panel supports the latitude given to the Secretary with respect to making awards after the peer review process has been completed. The standards allow the Secretary to consider a wide array of information in making awards, including "an applicant's ranking; recommendations of the peer reviewers

with regard to funding or not funding; information concerning an applicant's performance and use of funds under a previous Federal award; amount of funds available for the competition; and any other information relevant to a priority or other statutory or regulatory requirement applicable to the selection of applications for new awards."

Findings and Recommendations

There are a number of areas in which the peer review process might be improved; in many cases this improvement would not require a change in the standards, but in the way they are implemented. This section presents the study findings and recommendations in eight key areas: (1) enhancing the match between applications and review expertise, (2) reducing reviewer workload, (3) bolstering the professional development of reviewers and applicants, (4) clarifying the standards, (5) modifying several of the review criteria and weightings, (6) eliminating standardization of scores, (7) providing better feedback to applicants, and (8) exploring the use of technology. The goal of the recommendations presented below is to improve the quality of the peer review process such that scientifically appropriate research is supported, and the public interest is protected in the selection of grantees, while at the same time reviewers have an opportunity to learn from the expertise of their peers on the review panels.

1. Enhancing the Match Between Applications and Reviewer Expertise

Findings

According to a paper prepared by Department of Education (ED) staff (Karp et al.,

1995), there are a number of constraints on OERI's ability to obtain adequate numbers of good reviewers, including reviewer unavailability (e.g., too busy), conflict of interest, and unwillingness to serve (e.g., takes too much time, too little remuneration, no professional advantages). Further constraints are introduced by the need to ensure ethnic, racial, gender, and geographic diversity; the difficulty of finding experts in specialty areas; the current once-a-year submission dates, which result in competition for reviewers among institutes; and a lack of adequate funds for reviewer costs. With regard to cost, OERI is permitted by law to spend only up to 1 percent of program funds (by program account) for reviews. Funds cannot be moved from one program account to another or from salary and expense accounts to program accounts to pay for reviewers.¹ Finally, OERI staff have had far less time than staff in other federal agencies to locate reviewers or constitute review panels. In the fiscal year (FY) 1997 field-initiated studies (FIS) grant competition, staff had only 3 weeks to find reviewers. At the National Science Foundation, in contrast, staff generally have 6 weeks to select reviewers for a given competition.

Analysis of the fit between reviewer backgrounds and competition subject areas indicates that some of the FIS panels included one or no reviewers with research expertise. Moreover, in our analysis we found a relationship between panel credentials and review quality. The three panels that appeared to have no reviewers with research training and experience in the subject area of the competition produced no good reviews according to our criteria; the majority were indifferent or poor. Conversely, for the three panels identified as

¹ Karp et al. (1995) provide the following example: Prior to the FY 1996 FIS competition, the program was funded at approximately \$1 million per year, and OERI routinely received over 300 proposals. As a result of the legal cap on review expenditures, only \$10,000 was available to pay for travel, per diem, and honoraria for reviewers to read the applications.

having strong researchers in the field of the competition, we found that the majority of the reviews were good, and none were poor. This was a far greater proportion of good reviews than was found for the other panels. This finding may be a function not only of qualified reviewers, but also of the give and take that occurs in panel discussion, which further enhances review quality. In addition to our analysis, comments from applicants and reviewers indicate concern about the lack of research methodology and design expertise among reviewers.

Although the evaluations are to be accompanied by "concise written comments based on the reviewer's analysis of the strengths and weaknesses of the application with respect to each of the applicable evaluation criteria," many of the reviews were found to be cursory and descriptive. The reviewers had not provided analytic comments drawing on their background knowledge and expertise.

Recommendations

1-1. All reviewers should meet the requirements set forth by the standards.

The importance of reviewers' research knowledge and background is directly acknowledged in the third reviewer qualification cited earlier--"in-depth knowledge of theoretical perspectives or methodological approaches in the subject area of the competition." The standards appear to be an effort to raise the bar, ensuring a high-quality group of peer reviewers who will be able to lend expertise, insight, and stature to OERI's peer review efforts. By requiring that all members of a panel meet all three reviewer qualifications, the standards clearly indicate that creating a panel is not to be a "mix-and-match" effort, with one reviewer representing the research community, one practitioners, and one methods or theory.

There is little doubt that it takes considerable effort to find individuals who meet all

three qualifications in the standards. Moreover, some high-quality reviewers might be excluded from the process because they meet some but not all of the criteria. If OERI decided to change the standards so that each reviewer need not meet all three criteria, a two-tier process might be put in place. Such a process would entail the review of applications for their technical merit by reviewers with the requisite expertise during the first-tier review.

Applications that were technically sound would go forward to a second panel that would rate the applications on additional criteria. The National Institutes of Health (NIH) uses a two-tier approach in which the first panel judges applications on technical merit, while the second considers both technical merit (as judged by the first-tier review) and the relevance of the proposed study to the institute's programs and priorities.

1-2. OERI should consider establishing standing peer review panels within each institute.

Each institute would have a standing panel comprised of 25 to 30 individuals having the qualifications set forth in the standards. To comply with the standards, it would be important to include practitioners and policymakers with a solid understanding of research, and researchers with an understanding of policy and practice. Examples of practitioners or policymakers with research expertise are directors of assessment within state education departments and principals of laboratory schools who conduct research. Examples of researchers with knowledge of policy or practice are policy researchers, as well as researchers who work collaboratively with practitioners and policymakers.

Rules for establishing these standing panels would need to be developed, building on

the current rules for ad hoc panels. These rules might require additional regulations.² Included in these rules would be the process by which panelists are selected; the length of appointments (e.g., staggered terms of specified duration); and diversity criteria (e.g., ethnic background, gender, geographic location, urbanicity) that are not identified in the OERI standards, but are applied by other agencies. To ensure both high quality and prestige, standing panelists might be appointed by the OERI Assistant Secretary in consultation with the National Educational Research Policy and Priorities Board (NERPPB).

There are several important potential benefits of using a standing panel to review grant and cooperative agreement applications. First, doing so would increase the likelihood of having highly qualified reviewers for the competitions, including reviewers from underrepresented groups, because standing panels would afford greater professional satisfaction and prestige than is offered by the present ad hoc system. Furthermore, because standing panelists would know well in advance that they would be serving on specified dates, their availability to review applications would be ensured. Moreover, when standing panels made recommendations to an applicant about how to improve an application, there would be a good likelihood that many of the same reviewers would read a resubmission. (Multiple submission dates, which would also facilitate resubmission, are discussed below.)

In addition, if standing panels brought together the "best and the brightest," many people would be willing to serve because they would find the experience both educational and professionally satisfying. Providing panelists with letters of commendation from the Secretary of Education and remuneration commensurate with the task would offer further incentives,

² Lawyers from the Office of the General Counsel in the Department of Education noted this possibility (personal communication, September 1998).

although increasing the amount of reviewer compensation might require a change in legislation (because, as noted above, reviewer costs can currently total no more than 1 percent of the program budget).³ However, in our review of other agencies, we found remuneration to be at a level similar to that in the Department of Education.

One potential drawback of a standing panel is that it could be difficult to match the professional expertise of panel members with applications covering diverse topics and employing different qualitative and quantitative methodologies. This problem is particularly relevant because the FIS programs, especially in some institutes, attract such a diversity of applications. The problem might be solved by employing multiple subpanels with different areas of expertise or recruiting additional ad hoc reviewers should specific expertise be needed. At NIH, for example, the Division of Research Grants maintains a cadre of experts who are enlisted to review individual applications when additional expertise is needed. Standing panel members or staff initiate the request for these additional reviews, whose results are presented at the panel meeting.

Another potential problem with the use of standing panels for OERI is that the FIS competitions in some institutes receive 200 applications per year. If standing panels were to meet on an annual basis, reviewers could find themselves with too many applications to read at one time. There are several ways, however, to reduce the number of applications any given reviewer has to read. These include providing detailed evaluations for competitive applications only; using preliminary applications; assigning primary, secondary, and tertiary reviewers to

³ If the standing panel performed a variety of roles of which peer review of applications was only one, panelists could also be employed as consultants under program or salaries and expenses authority.

applications; and increasing the number of submission dates and the number of times panels meet each year to two or three. All of these options are discussed in a subsequent section.

It might also be difficult to recruit standing panelists to serve for several consecutive years given that during that time, they might not be able to apply for funds from the institute for which they were serving as a reviewer. Procedures might be put in place so that panelists could submit applications in the general content area of the institute for which they were serving as a reviewer.⁴ At NIH for example, applications of standing panel members are reviewed by an ad hoc panel convened for the purpose. If a two-tier review process were used, panelists who had submitted applications might be able to recuse themselves from the second-tier review if their applications were being considered. They could be replaced by a member of their three-member panel from the first tier. In the case of center reviews, those who had submitted an application would not participate on the review panel.

1-3. The quality of the research design should be rated only by reviewers with appropriate technical expertise.

To enhance the quality of the reviews, additional reviewers with expertise in the research design associated with specific applications could evaluate the applications on this criterion. This approach might require the use of ad hoc reviewers with specialized expertise. A two-tier approach as previously described, employing first-tier reviewers with the requisite technical expertise, would also support this recommendation.

1-4. The size of review panels should be increased.

It is important to have a panel of a sufficient size, along with a concomitant increase in

⁴ This would require consultation with Department of Education lawyers.

the number of applications read by a given panel, so that there will be a high probability of the panel's reviewing applications that range in quality (from those likely to be successful to those likely to fail), sufficient breadth of perspective and expertise on the panel to ensure that innovative applications will not be overlooked because reviewers fail to understand them, and an opportunity for individual panelists to calibrate their reviews/ratings against applications they have not read quite so carefully before formulating their independent ratings. The National Science Foundation (NSF) accomplishes this by dividing a comparable number of applications among a smaller number of larger panels; a comparable number of applications is assigned for written reviews to each panel member, but all panelists have the opportunity to examine all applications to be reviewed by the panel. A two-tier process might accomplish some of the same objectives, but would require that first-tier panel summaries as well as ratings be provided to second-tier reviewers to give them the perspective afforded by access to additional applications.

1-5. The database of reviewers should be improved.

Because many OERI staff are not researchers in the fields in which they manage research, it would be valuable for them to undertake reviews of bibliographies in those fields, if they do not already do so, in order to locate reviewers with the qualifications specified by the standards. Staff should also systematically solicit names from professional association representatives, grantees, and panelists on ad hoc or standing review panels (if such standing panels are created as recommended above). Many other agencies maintain a central database of reviewers that staff can access. If standing panels are formed within each institute, a database of reviewer names will be useful for selecting additional specialized reviewers.

The identification of standing panelists will require a more elaborate selection process. In all cases, staff might systematically evaluate the performance of reviewers and continue to use only those reviewers that demonstrated satisfactory performance, as evidenced by mature judgment, balanced perspective, objectivity, ability to work in a group, reliability, and integrity as well as the preparation of adequate review comments. Reviewers who cancel at the last minute, come to peer review meetings unprepared, and write minimal comments have a negative impact on the review process and must be disqualified from participating in future reviews.

1-6. OERI staff should attempt to issue grant announcements earlier in the fiscal year, thereby increasing the amount of time available for selecting and assigning reviewers.

As noted earlier, staff at other agencies have more time than OERI staff to select reviewers. If standing panels were used, staff would not have to select reviewers, but would have to assign applications to the appropriate reviewers (at least three per application) and, if appropriate, assign lead and secondary reviewers. Additional time would also give staff a chance to assess the content of the applications prior to finalizing review teams so as to optimally match reviewer expertise to application topics, as well as to select ad hoc reviewers if necessary.

1-7. There should be well-established submission dates, staggered by institute.

In some previous competitions, reviewers have not had enough time to review applications because there has been such a small time window between the submission of applications and the conduct of the review. Well-established submission dates for each

institute, with the dates for the different institutes staggered over a 3- to 4-month period, would ensure the submission of applications in time for staff to select reviewers carefully and constitute appropriate review panels. In addition, the use of staggered submission dates would make it possible for reviewers to read applications for more than one institute. In making grant announcements, staff could estimate the amount of funding available for a competition, using the previous year's funding as a floor.

2. Reducing Reviewer Workload

Findings

As mentioned previously, many of the reviews were found to be cursory and descriptive rather than comprehensive and analytic. According to reviewers and Department of Education staff, this is due in part to the short time available for preparing reviews, and in part to the large volume of material reviewers have to review.

Recommendations

2-1. Logistics and other support for reviewers should be increased

Reviews could be improved by giving reviewers more time for their reviews, both prior to on-site meetings (as discussed above) and during the panel sessions. Making computers and disks available during the panel sessions for all reviewers who wanted them would also help reviewers provide more coherent and elaborate comments (see the discussion of the use of technology later in this chapter).⁵

2-2. Applications that are non research should be disqualified prior to peer review.

⁵ ED staff do not believe implementing this recommendation is feasible at the current time, although panelists we interviewed requested that such an option be explored.

The definition of research is very broad, and might be difficult to narrow given congressional support for a broad definition.⁶ Nonetheless, 5 or 6 of the 29 FY 1997 FIS applications reviewed were not really research studies. Most of these applications were from practitioners who were seeking funds to develop and implement a program, curriculum, or software. For example, one applicant proposed implementing a variety of after-school activities for students, while another proposed constructing an interactive website.

Staff should work with a common definition of research. They should review applications to identify those that do not fit the definition of research. Providing staff with a screening checklist might help weed out nonresearch applications. If the staff decide to submit such applications for peer review anyway (for fear of a challenge), they should give reviewers more explicit assistance in responding to these applications. Reviewers should be encouraged to explain in their reviews how the applicant could revise the application to focus it on research (if that appears feasible) or improve its development or implementation proposal (or both).

2-3. Detailed evaluations should be provided for competitive applications only.

There are currently no procedures for screening applications for merit prior to the reviewers' full evaluation. Reviewers might initially screen each application submitted under a particular competition as likely to be "competitive" or "noncompetitive." Only applications rated as likely to be competitive by a majority of reviewers would then receive a detailed

⁶ Public Law 103-227, the Educational Research, Development, Dissemination, and Improvement Act of 1994, Section 912(i)(6), defines educational research: "The term educational research includes basic and applied research, inquiry with the purpose of applying tested knowledge gained to specific educational settings and problems, development, planning, surveys, assessments, evaluations, investigations, experiments, and demonstrations in the field of education and other fields relating to education." Section 912 (i)(7) defines the term field-initiated research: "The term field-initiated research means education research in which topics and methods of study are generated by investigators, including teachers and other practitioners, not by the source of funding."

evaluation. For those deemed noncompetitive, staff could prepare a summary of the discussion and send it to the applicant, with extended suggestions for improvement. An advantage of this procedure is that it would increase the time available to peer reviewers for evaluating high-quality applications and shift more of the routine work to staff. If this process were to be used, staff would have to define cut-off scores for noncompetitive applications.

2-4. Decrease the number of full applications through the use of preliminary applications.⁷

A pre-application is an abbreviated grant application, which is typically judged on a subset of evaluation criteria (management is not described, for example). Pre-applications are reviewed by staff and/or peer reviewers. Those applicants whose pre-applications are judged favorably are then encouraged to submit full applications, while others are discouraged (although anyone is free to submit).⁸ Many agencies use such preliminary applications, including the Fund for the Improvement of Postsecondary Education (where peer reviewers read pre-applications) and NSF (where they are read by program staff and/or peer reviewers). Use of pre-applications would reduce the number of full applications received; it would also allow staff to estimate the number of full applications to be received, and thus the types and numbers of peer reviewers needed. Drawbacks to the use of pre-applications include reduced information available to reviewers for forming their initial judgments and possible lengthening of the review process (which would now include two stages of application review).

2-5. The number of pages permitted for center applications should be decreased

⁷ A change in the regulations might be required to implement this procedure.
⁸ The current standards for peer review do not permit pre-applications.

and for center application attachments limited.

This would be one way to reduce what has proven to be a heavy reviewer workload. At present, some center applications are many hundreds of pages long.

2-6. Provide planning grants for center competitions.

If operationalized, initially, applicants for center funds would submit a preliminary application. Page limits would be specified. The focus would be on a conceptual framework, rather than on individual studies. Selected applicants would then be provided planning grants to prepare full applications that would provide more detail on the individual studies to be undertaken, as well as on other criteria, such as the management plan. As at NSF, a panel would be convened to rate applications at each stage of the review.

2-7. Primary, secondary, and tertiary reviewers to applications should be assigned⁹

This type of review is characterized by larger review panels (with anywhere from 8 to 25 reviewers); thus this recommendation would apply only if the size of the review panels were increased. To decrease reviewer workload, OERI might institute a system in which different reviewers would have differing responsibilities for evaluating the same applications. As with the procedures used at NIH, some reviewers might read applications more thoroughly and have primary responsibility for presenting findings and recommendations to the review panel. All panel members could read an application if they chose to do so, and could comment on and ask questions about that application in the discussion. A potential disadvantage of this approach is that primary reviewers might dominate the panel discussion.

⁹ This recommendation might require a change in the regulations.

2-8. Establish multiple submission dates each year.

This approach is used, for example, at NIH, where study panels meet several times a year, thereby reducing the number of applications read at any given time. A disadvantage of this approach is that it would engender additional costs and be time-consuming for staff. In addition, several OERI staff believe multiple submission dates might actually increase the total number of applications submitted each year because applicants would have more than one opportunity to submit an application.

3. Bolstering Professional Development

Findings

According to Department of Education staff and reviewers, having higher-quality applications would facilitate the review process. In addition, many reviewers expressed concern that some applicants with worthy projects do not have the capability to prepare adequate applications.

Although the standards emphasize independent judgments in evaluation, comments from reviewers during our interviews, as well as their comments written on evaluation forms, indicate that this is not always the case. Instead, they revise their scores and written comments during the panel discussion of application strengths and weaknesses. Thus decisions about final scores, as well as whether an application is highly recommended, recommended, or not recommended, are sometimes made collectively. Moreover, across years and types of competition, many applicants we interviewed assessed reviewer comments as not very useful or comprehensive.

There has been considerable confusion about the appropriate qualifications of each reviewer on the part of OERI staff. Some OERI staff expressed the view that at least one of the peer reviewers of each FIS application—but not all three—should be a researcher in the subject area of the competition. Staff indicated that panels have been constituted by selecting one researcher with subject matter expertise, one methodologist, and one person with expertise in policy or practice in the area of the competition which is counter to the standards as we interpret them.

Recommendations

3-1. Enhance training for applicants.

OERI staff currently offer some technical assistance to interested potential applicants at regional and national professional association meetings. These efforts should be increased and made more systematic. In other federal agencies, it is not uncommon for project or program officers to provide detailed advice to potential applicants. Staff of the federally funded regional laboratories, comprehensive centers, and research centers could also provide such assistance.

3-2. Provide more in-depth training and support for reviewers.

Peer reviewers currently receive some training with regard to evaluation criteria and other review procedures. For example, reviewers are provided information about their responsibilities as panel members and completing standard evaluation forms. This training is provided through the materials sent to panelists prior to their review and as part of the basic orientation at the beginning of the panel meeting.

OERI should consider providing more in-depth training to peer reviewers. Reviewers

need to understand the requirements for *independent* assessment, scoring, revision, and assignment of applications to funding categories. They also need to understand that their comments will be the only substantive information unsuccessful applicants receive; thus those comments should help applicants understand the decisions that were made, as well as improve their future submissions. Such training might include detailed instruction in properly evaluating grant applications. Examples of exemplary reviews might also be provided to illustrate the standards to which reviewers should hold themselves in rating applications.

Another means of improving reviewers' performance is to provide them with a set of questions to guide their reviews. Given that many of the applications fall into three or four broad categories of research design (e.g., curriculum or program development with pilot test or observation; evaluation of a curriculum or intervention; and implementation study of a state or local policy using a variety of methods, such as interviews, focus groups, classroom observation, and document analysis), it may be possible to draft a set of questions specifically appropriate to these broadly used designs. For example, there are certain questions one can ask about equivalence of treatment of control and comparison groups; about the documentation of implementation; and about the relationships among intervention, instruments, and various dependent variables. These questions could be suggested as ones reviewers might ask about the applications they review.

Moreover, many reviewers who write detailed and cogent comments on design simply do not appear to concern themselves with timelines, benchmarks, or budgets, possibly because they may not be the most appropriate judges in these areas and do not know the correct questions to ask. They may focus on small details (such as missing resumes for research

assistants) because they do not know what else they should be considering. This problem is most critical with respect to center reviews, where millions of dollars are at stake. OERI should provide a series of specific questions for reviewers to ask about budgets and management plans (and also some ideas about what items are off limits, such as entertainment expenses).

Providing reviewers with elaborated scoring rubrics that would serve as criteria against which to rate applications is another way to improve the reviews. Such rubrics would explain the meaning of different numerical scores, e.g., what it means to assign 29 of 30 or 20 of 30 points. The rubrics would differ by evaluation criterion.

3-3. Provide professional development for OERI staff to ensure that they understand the requirements of the standards.

Although many OERI staff are familiar with the requirements of the standards, some are not. For those staff needing more information (e.g., staff new to OERI), training would be useful and appropriate, and should be provided.

4. Clarifying the Standards

Findings

In addition to the individual qualifications for peer reviewers, the standards require that OERI (the Secretary) select "to the extent feasible...peer reviewers for each competition who represent a broad range of perspectives." This requirement is not further elaborated in the discussion of the standards, so it is not immediately apparent how it should be implemented.

With regard to conflict of interest, the standards specify that reviewers for grants and

cooperative agreements are considered employees of the Department, and as such they are subject to provisions of 18 U.S.C. 208, CFR 2635.502, and Department policies that implement these provisions. Under those rules, reviewers are considered to have a conflict of interest if they or their immediate family, a for-profit or nonprofit organization in which they serve, or any organization or person with whom they are negotiating or have an arrangement concerning prospective employment have a financial interest in the application they are reviewing. For FIS competitions, the Department asks for a waiver so reviewers can serve even if applications from their university are expected. However, no reviewer is ever assigned to read an application from his or her university or other employer.

Although reviewers do not read applications from their own university, our analysis indicated that some reviewers had had prior professional relationships with the applicants whose applications they were reviewing. While this may not constitute conflict of interest as defined by the standards, it could predispose reviewers to judge applicants by those prior relationships, rather than solely on the merit of the application. Moreover, the appearance of a conflict can constitute a serious problem.

Finally, while the text of the application package may point out that the priorities are nonbinding, it remains unclear how reviewers and others should view those priorities in determining application strengths/weaknesses, rankings, and awards. Some applicants said the priorities were critical in their decision to submit an application. Some reviewers assigned points for national significance if an application addressed these priorities.

Recommendations¹⁰

¹⁰ Clearly, these recommendations would require a change in the standards.

4-1. The term "multiple perspectives" in the standards should be elaborated.

Gender, race/ethnicity, and geographic considerations (including rural/urban perspectives) should be taken into account when review panels are constituted. The peer reviewer selection process might be made similar to the process at NIH or the Office of Special Education and Rehabilitative Services (OSERS) within the Department of Education. At NIH, factors such as geographic distribution and minority and female status must be considered in selecting review group members (National Institutes of Health, 1992:6). In OSERS, staff consider the overall representativeness of the panels convened for a competition, especially the presence of underrepresented groups, such as minority groups and persons with disabilities. Panels should also be constituted to take into account diverse disciplinary background, theoretical orientation, methodological approach, and research role.

4-2. Conflict of interest should be further defined to include professional relationships.

OERI might model its procedures on those used at NSF. NSF informs reviewers that "they may not participate in the review of any proposal in which they or a member of their immediate family or an organization of which they are or may become a part has a financial interest, nor may they be in the room when such a proposal is discussed." In addition, in instructions to reviewers, NSF provides examples of conflict of interest that go beyond financial interest. These include "reviewer would be directly involved in the project, e.g., as a consultant or collaborator; reviewer is from the same institution as the proposer; reviewer and proposer have been related recently as a student and thesis advisor or postdoctoral advisor; reviewer and proposer are known to be close friends or open antagonists; reviewer and

proposer have collaborated recently on a related project; reviewer and proposer were co-authors on a paper published in the last four years."

4-3. Do not list priorities for FIS competitions.

Panelists interviewed supported the use of priorities for directed research competitions, but not for FIS competitions. Other agencies with more funds available for research run several kinds of competitions. Research priorities are not specified for FIS grant competitions, whereas they are specified for directed research grant competitions.

5. Modifying the Review Criteria and Weightings

Findings

The standards specify broad evaluation criteria, and specific factors under each, from which OERI may select review criteria for each grant or cooperative agreement competition. The broad criteria include national significance, quality of the project design, quality and potential contributions of personnel, adequacy of resources, and quality of the management plan. The factors offer a wide range of options, with some oriented to research competitions and others more appropriate to demonstrations or program grants. The rules allow for complete discretion with respect to which of the broad categories is used in any competition, as well as which (if any) of the specific factors are selected.

Applicants and reviewers for both FIS and center competitions indicated confusion about the meaning of the national significance criterion and how it should be addressed. In our assessment of the reviews, we found that reviewers and applicants often define national significance as the importance of the problem to be addressed. They do not interpret it to

include the potential contribution of the project to the development and advancement of theory and knowledge in the field in addressing an important problem.

As noted, the standards allow for considerable flexibility in both the broad evaluation criteria and specific factors that are selected for each competition. Yet the broad criteria are the same for the FIS and center competitions. This is problematic because there is such a difference in scale between the FIS and center competitions. For example, the evaluation criteria for quality of project design for the center competitions are currently written as if reviewers were evaluating a single research study (as is the case with the FIS competition), rather than a series of studies, which is what most center applications propose. Partly as a consequence of this, we found that reviews of center applications rarely discuss and assess the quality of the proposed study design. Instead, they focus on the overall conceptual framework. As a second example, reviewers and Department of Education staff commented on the need for increasing the weight given to the management criterion for centers. The management component is very important for centers because most centers are located in multiple, geographically diverse sites and involve the management of complex research activities across these sites. This is generally not the case for FIS.

Recommendations

5-1. Clarify the meaning of "national significance."

The review criterion of national significance should be clarified so reviewers and applicants understand that it refers to both the importance of the problem to be addressed and the potential contribution of the project to the development and advancement of theory and knowledge in the field.

5-2. Further elaborate the project design criterion for center applicants.

As noted, reviews of center applications often do not address the individual studies being proposed. In part, this may be due to a lack of clarity in the evaluation criteria about whether such analysis is needed.¹¹

5-3. Increase the weighting for management for center applications.

Different weightings for FIS and center competitions on this criterion may be warranted given the complex activities across sites typically undertaken by Centers.

6. Eliminating Standardization of Scores

Findings

With respect to standardization of scores, some grant competitions and some institutes have elected to use an Department of Education standardization process designed to correct for possible bias introduced by different reviewers' approaches to assigning raw scores.

According to a report prepared for the U.S. Department of Education by Advanced Computer Systems, Inc. (1992:10), the standardization process is based on a set of assumptions about the distribution of applications:

...that the varying quality of applications in the entire pool of applications is normally distributed, i.e., a similar number of good, average, and poor applications is reviewed by each panel; the applications distributed to each panel are normally distributed; all panels have the same training and direction; and any resulting deviation is due to reader

¹¹ Implementing this recommendation might necessitate changes in the standards.

bias (Analysis of the Grants and Contracts Management System Score Standardization Program

Staff we interviewed indicated that FY 1997 FIS applications were not randomly assigned to panels within institutes, but assigned on the basis of application topic. Moreover, interviews with reviewers revealed that some were aware of the standardization process and had adjusted their scores to ensure that favorite applications would have a better chance of being funded; other reviewers were not aware of the process, and thus had rated applications without regard to standardization. In examining applications, we found that at least five of those sampled were not research, and the scores for these applications may have skewed the distribution.

Recommendation

6-1. Do not employ standardization for FIS and center competitions.

Standardization would be rendered unnecessary by the use of second-tier panels that would review and rank the applications of finalists from the first tier (as discussed earlier) or standing panels that would review at one time all applications submitted to an institute.

7. Providing Feedback to Unsuccessful Applicants

Findings

OERI staff currently send copies of the reviewers' Technical Review Forms to each unsuccessful applicant. These forms contain the applicants' raw scores, along with reviewers' descriptions of strengths and weaknesses associated with each evaluation criterion and a brief

summary of the review. Many applicants reported that the reviews they received were not useful. As previously mentioned, our analysis indicated that many of the reviews were cursory.

Recommendation

7-1. OERI staff should consider providing unsuccessful applicants with more detailed feedback on their applications.

Previous recommendations aimed at improving the review process should help achieve this objective. In addition, staff might provide a written summary of the panel discussion of each application, a procedure used at both NSF and NIH (where it constitutes the main feedback to applicants and incorporates written review comments).

8. Exploring the Use of Technology

Findings

Our review of the peer review process in other agencies revealed little evidence of the use of technology in peer review (although NSF uses a method for electronic filing of applications that enables applicants to track the progress of their application). Nonetheless, both the OERI staff we interviewed and expert panel members expressed interest in pursuing the use of technology to expedite the peer review process.

Recommendation

8-1. OERI should consider a small pilot project to determine whether and how technology could be used to support the peer review process.

For example, if a two-tier review model were adopted, reviewers might read

applications during the first tier and download their reviews onto a website devoted to the competition. After all reviewers had had the opportunity to read each other's reviews, they might conference via telephone or e-mail to resolve differences and decide which applications should move forward.

Summary of Recommendations

This section highlights the study recommendations that are most central to improving the OERI peer review process.

First, standing panels of 25 to 30 reviewers should be established in each institute. Reviewers should be carefully selected to ensure that each meets the criteria established by the standards. Panels should be constituted to ensure ethnic, racial, geographic, and gender diversity. Moreover, a balance between senior and junior scholars should be sought. Proposed panelist slates should be approved by the institute directors and the Assistant Secretary for OERI, with consultation from the NERPPB.

The reviewers on these standing panels should serve set (e.g., staggered 3-year) terms and form the core of reviewers for each institute. For the center competitions, a subset of standing panelists should be used. Decisions about which panelists to select for a center competition and the number needed should be based on the applications received for a particular competition. The subpanelists could also serve as midterm reviewers, thus ensuring consistency in the review process.

For field-initiated studies (FIS) competitions, there are two options for the review process. The first would entail the formation of six to eight-member subpanels from the

membership of the standing panel; these subpanels would provide the first tier of review. The first-tier review process would function much like the current process, except the subpanels would comprise primarily standing panelists and would be expanded from three members to six to eight members to provide a broader context for the review. Applications would be allocated to subpanels on the basis of the panelists' subject area expertise and experience. If the review of some applications required special technical expertise, the subpanels could be supplemented with ad hoc reviewers. During each review cycle, the team leaders of each subpanel would meet an additional day for a second-tier review to rate all the top-ranked applications from the first-tier subpanels.

A second option for the FIS review process resembles the process used at the National Institutes of Health, where the entire panel reads all applications. At NIH it is typical for a group to review 75 to 100 applications at each meeting. Each member is asked to prepare detailed reviews for a dozen or more applications. The meetings are conducted by a chair who is a peer, assisted by a staff member. Those preparing the written reviews lead the discussion of the applications assigned to them. Each application is discussed and considered. Decisions not to recommend for further consideration are made by majority vote.¹² If a member disagrees, he or she can submit a minority report, and when there are two or more dissenting members, a minority report must be drafted. Members who cannot assess the merits of a proposal can abstain from voting, although abstentions are not encouraged. Applications can also be deferred (perhaps for a site visit or to obtain additional information). Those applications not rejected or deferred are assigned a priority score by each member. These

¹² This procedure is not permissible under the current standards.

scores are averaged by the staff member after the meeting. In addition, a summary statement for each application, prepared by the staff person involved in the review for transmittal to the council and the applicant, shows a percentile ranking for the application against a reference base of all applications reviewed by the committee over three meetings, including applications not recommended for funding or deferred (helping to minimize the effects of a single meeting). The written comments of panel members and the panel discussions are the basis for these summary statements.

Standing panels might also be involved in other activities, such as recommending how OERI could help foster research in a particular area in which good applications had not been received; helping to select new panelists and ad hoc reviewers; and reviewing grant-produced products, especially once the Phase 3 standards have been put in place. Panels might also provide continuity in the assessment of applications so that rejected applications that had been revised and resubmitted would be reviewed by at least some of the same people. In addition, panelists could serve on midterm review teams for operating centers.

Given that some institutes receive up to 200 applications annually, methods for reducing reviewer workload should also be considered. Several possibilities are elaborated in this report. They include, for example, the use of preliminary reviews to reduce the number of full applications receiving a detailed evaluation, and the use of pre-applications, with only a subset of applicants being asked to prepare a full application.

This report also makes other recommendations for improving the OERI peer review process. First, professional development should be enhanced for Department of Education staff, especially those new to the process, as well as for applicants and reviewers.

Additionally, reviewers would benefit from questions to guide their reviews and from elaborated scoring rubrics. The standards would benefit from clarification in several areas: the term "multiple perspectives" should be further defined to ensure that panel membership is balanced by gender, race/ethnicity, and geographic location as well as by disciplinary background, theoretical orientation, methodological approach, and research role; conflict of interest should be defined to include professional relationships as a source of conflict; and priorities should not be listed for FIS competitions. Modifying the review criteria and weightings would also enhance the process: the meaning of "national significance" should be clarified; for center competitions, the project design criterion should be elaborated; and weighting for management should be increased. Standardization of scores should be eliminated as well; the use of second-tier panels and standing panels would make this process unnecessary. Finally, more detailed feedback should be provided to unsuccessful applicants, and the use of technology in the peer review process should be explored.

RESOURCE MATERIAL

American Educational Research Association

Ethical standards of the American Educational Research Association.

Educational Researcher 21(7):23-26.

1998 Research Policy Notes. OIA Info Memo, January/February 1998.

Washington, DC: American Educational Research Association.

Barkdoll, G.L.

1995 Standards and their use in the Food and Drug Administration (FDA).

Commissioned paper for the Office of Educational Research and Improvement, U.S. Department of Education.

Chubin, D.E.

1986 Much ado about peer review. *BioScience* 36(1):18-21.

1994 Grants peer review in theory and practice. *Evaluation Review* 16(1):20-30.

Joseloff, S. H.

1984 The Making of NIH Research Grants, *NIH Record*, August 14 and August 28. (Updated by Andrea Shreve, May, 1992.) Bethesda, MD: National Institutes of Health.

Justiz, M.J., and H.N. Moorman

1985 New NIE peer review procedures. *Educational Researcher* January:5-11.

Kostoff, R.N.

1994a Assessing research impact. Federal peer review practices. *Evaluation Review* 18(1): 31-40.

1994b Assessing research impact. Semiquantitative methods. *Evaluation Review* 18(1):11-19.

Kostoff, R.N., H.A. Averch, D.E. Chubin

1994 Research impact assessment. Introduction and overview. *Evaluation Review* 18 (1):3-10.

McInerney, M., D. Levin, M. Kane

- 1996 *Review Activities for Implementing Part E Program Agenda: Review of Selected Literature on Peer Review Procedures.* Washington, DC: Chesapeake Institute.

National Endowment for the Humanities

- 1995 *1995 Annual Report.* Washington, DC: National Endowment for the Humanities.
- No *Educational Development and Demonstration.* Grant Application
- Date Instructions and Forms. Washington, DC: National Endowment for the Humanities.

National Institutes of Health

- Orientation Handbook for Members of Scientific Review Groups, Interim Revision (March 1992).* Bethesda, MD: National Institutes of Health.
- No *Peer Review of NIH Research Grant Applications.* (Prepared from slides)
- Date Bethesda, MD: National Institutes of Health

National Research Council

- Improving Research through Peer Review.* Prepared by the Committee on Peer Review Procedures, Board on Agriculture. Washington, DC: National Academy Press.

National Science Foundation

- Grant Policy Manual (NSF 95-26).* Arlington, VA: National Science Foundation.
- 1997 *Grant Proposal Guide (October)* Arlington, VA: National Science Foundation.
- 1998 Personal interviews conducted by Lana Muraskin with staff members of the National Science Foundation, Arlington, VA.

Noble, J.H., Jr.

- Peer review: Quality control of applied social research. *Science* 185 (September 13, 1974):916-921.

Office of Educational Research and Development

No New Application for the Field Initiated Studies Educational Research

Date Grant Program. CFDA Numbers: 84.305F, 306F, 307F, 308F, 309F.

Applications due January 5, 1996. Washington, DC: U.S. Department of Education.

Office of Educational Research and Improvement

1995a Application for a Grant under the National Educational Research and Development Center Program, Closing Date: December 15, 1995. CFDA Numbers: 84.305A, 84.305B, 84.306A, 84.307A, 84.308A, 84.309A, 84.309B. Closing date: December 15, 1995. Washington, DC: U.S. Department of Education.

1995b 1995-96 Technical Review Plan for the National Educational Research and Development Center Program. CFDA Numbers: 84.305A, 84.305B, 84.306A, 84.307A, 84.308A, 84.309A, 84.309B. Washington, DC: U.S. Department of Education.

1995c Reinventing PEER Review in OERI. Paper transmitted September 22, 1995 by note to members of the Board, National Institute on Early Childhood Development and Education, OERI, U.S. Department of Education.

1995d Standards for the Conduct and Evaluation of Activities Carried out by the Office of Educational Research and Improvement (OERI)—Evaluation of Applications for Grants and Cooperative Agreements and Proposals for Contracts (34 CFR Part 700). Washington, DC: U.S. Department of Education.

1996a Application Technical Review Plan for the Field-Initiated Studies Research Grant Program under the National Institute on the Education of At-Risk Students. Washington, DC: U.S. Department of Education

1996b Application Technical Review Plan, Field-Initiated Studies Research Grant Program, National Institute on Early Childhood Development and Education. Washington, DC: U.S. Department of Education.

- 1996c Technical Review Plan, Field-Initiated Studies Research Grant Program, National Institute on Student Achievement, Curriculum, and Assessment, CFDA Number 84.305F (Amended). Washington, DC: U.S. Department of Education.
- 1996d 1996 Evaluation Plan, Field-Initiated Studies Program, National Institute on Educational Governance, Finance, Policymaking and Management. CFDA Number 84.308F. Washington, DC: U.S. Department of Education.
- 1996e 1996 Evaluation Plan, Field-Initiated Studies Program, National Institute on Postsecondary Education, Libraries, and Lifelong Learning (PLLI), OERI. CFDA Number 84-309F (Revised July, 1996). Washington, DC: U.S. Department of Education.
- 1997a Application for a Grant under the National Educational Research and Development Center Program. CFDA Number: 84.305R. Closing Date: May 28, 1997. Washington, DC: U.S. Department of Education.
- 1997b Application for a Grant under the National Educational Research and Development Center Program. CFDA Number: 84.308B. Closing Date: June 6, 1997. Washington, DC: U.S. Department of Education.
- 1997c Application Technical Review Plan for the FY 1997 Field-Initiated Studies Research Grant Program under all National Research Institutes. Washington, DC: U.S. Department of Education.
- 1997d Combined Application for the Field-Initiated Studies Educational Research Grant Program, Serving All National Research Institutes. CFDA Numbers: 84.305F, 84.306F, 84.307F, 84.308F, 84.309F. Applications Due: June 9, 1997. Washington, DC: U.S. Department of Education.
- 1997e Technical Review Plan for the National Research Center on Early Learning, FY 1997. CFDA Number: 84.305R. Washington, DC: U.S. Department of Education.

- 1997f Technical Review Plan for the National Research Center on Early Reading, FY 1997. CFDA Number: 84.305R. Washington, DC: U.S. Department of Education.
- 1997g Technical Review Plan for the National Research Center on Policy and Teaching Excellence, FY 1997. CFDA Number: 84.308B. Washington, DC: U.S. Department of Education.
- 1997h *Technology Innovation Challenge Grants*. Washington, DC: U.S. Department of Education.

Office of Postsecondary Education

- 1997 The Comprehensive Program: Information and Application Materials. Prepared by the Fund for the Improvement of Postsecondary Education. Washington, DC: U.S. Department of Education.

Office of Technology Assessment

- Federally Funded Research: Decisions for a Decade*. OTA-SE490. Washington, DC: U.S. Government Printing Office.

Public Health Service

- Application for a Public Health Service Grant. *PHS 398*, Rev. 5/95.

Sanders, H.J.

- Peer review: How well is it working? *C&EN* March 15, 1982:32-43.

Shulman, L.S.

- 1985 Peer review: The many sides of virtue. *Educational Researcher* 14(1):12-13.

U.S. Department of Education

- 1992a *Analysis of the Grants and Contracts Management System Score Standardization Program*, Revised February 6, 1992. Washington, DC: U.S. Department of Education.
- 1992b Discretionary Grant Planning, Review, and Award Procedures. Administrative Communications System Transmittal Sheet 92-39, dated June 24, 1992. Grants and Contracts Service. Washington, DC: U.S. Department of Education.

- 1997a *A Basic Guide to the Discretionary Grants Process of the U.S. Department of Education*. Office of the Chief Financial Officer, Grants Policy and Oversight Staff. Washington, DC: U.S. Department of Education.
- 1997b Combined Application for the Field-Initiated Studies Educational Research Grant Program. Available from the Superintendent of Documents, U.S. Government Printing Office (no. 1997-425-342/70226). Washington, DC: U.S. Department of Education.
- 1997c The Comprehensive Program: Information and Application Materials, Fiscal Year 1998. Washington, DC: U.S. Department of Education.
- 1997d Education Department General Administrative Regulations, 34 CFR Parts 74, 75, 76, 77, 79, 80, 81, 82, 85, and 86, March 6, 1997. Available from the Superintendent of Documents, U.S. Government Printing Office (no. 1997-426-871/70347). Washington, DC: U.S. Department of Education.

U.S. Congress, House

Public Law 103-227, 103d Congress March 31, 1994. Goals 2000: Educate America Act. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office

Peer Review: Reforms Needed to Ensure Fairness in Federal Agency Grant Selection. Washington, DC: U.S. General Accounting Office

U.S. Government Printing Office

1995 Part II. 34 CFR Part 700. Standards for the Conduct and Evaluation of Activities Carried Out by the Office of Educational Research and Improvement(OERI)-Evaluation of Applications for Grants and Cooperative Agreements and Proposals for Contracts; Rule. *Federal Register* 60(178) September 14, 1995:47808-47813.

Appendix A
Standards for Peer Review



**U.S. Department of Education
Office of Educational Research
and Improvement**

**Standards for the Conduct
and Evaluation of Activities
Carried Out by the Office
of Educational Research
and Improvement
(OERI)—Evaluation of
Applications for Grants and
Cooperative Agreements and
Proposals for Contracts
(34 CFR Part 700)**

September 1995

U. S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Educational Research Policy and Priorities Board

On March 31, 1994, President Clinton signed Public Law 103-227, which includes Title IX — the "Educational Research, Development, Dissemination, and Improvement Act of 1994" (the "Act"). The Act restructured the Office of Educational Research and Improvement (OERI) and endowed it with a broad mandate to conduct an array of research, development, dissemination, and improvement activities aimed at strengthening the education of all students. The Act also required the establishment of a National Educational Research Policy and Priorities Board (the "Board") to work collaboratively with the Assistant Secretary to identify priorities to guide the work of OERI.

The legislation directed the Assistant Secretary to develop, in consultation with the Board, such standards as may be necessary to govern the conduct and evaluation of all research, development, and dissemination activities carried out by the Office to ensure that such activities meet the highest standards of professional excellence. The legislation required that the standards be developed in three phases. These regulations implement the first phase of the standards. The Assistant Secretary will publish at a later date additional proposed regulations to implement the remaining standards in accordance with the timelines established in the Act. The legislation requires the Board to review and approve the final standards.

FOR FURTHER INFORMATION CONTACT: Edward J. Fuentes, U.S. Department of Education, 555 New Jersey Avenue, N.W., Room 600, Washington, D.C. 20208-5530. Telephone (202) 219-1895. Internet electronic mail address: stan_questions@inet.ed.gov

The list of subjects in 34 CFP Part 700 are as follows:

Education, Educational research, Elementary and secondary education, Government contracts, Grant programs-education, Libraries, Reporting and recordkeeping requirements.

The Secretary amends Chapter VII of Title 34 of the Code of Federal Regulations by adding a new Part 700 to read as follows:

Part 700—Standards for the Conduct and Evaluation of Activities Carried Out by the Office of Educational Research and Improvement (OERI)—Evaluation of Applications for Grants and Cooperative Agreements and Proposals for Contracts.

Subpart A—General

Sec.

- 700.1 What is the purpose of these standards?
- 700.2 What activities must be governed by these standards?
- 700.3 What additional activities may be governed by these standards?
- 700.4 What definitions apply?
- 700.5 What are the processes of open competition?

Subpart B—Selection of Peer Reviewers

- 700.10 When is the peer review process used?
- 700.11 Who may serve as peer reviewers?
- 700.12 What constitutes a conflict of interest for grants and cooperative agreements?
- 700.13 What constitutes a conflict of interest for contracts?

Subpart C—The Peer Review Process

- 700.20 How many peer reviewers will be used?
- 700.21 How are applications for grants and cooperative agreements evaluated?
- 700.22 How are proposals for contracts evaluated?

Subpart D—Evaluation Criteria

- 700.30 What evaluation criteria are used for grants and cooperative agreements?
- 700.31 What additional evaluation criteria shall be used for grants and cooperative agreements?
- 700.32 What evaluation criteria shall be used for contracts?

Subpart E—Selection for Award

- 700.40 How are grant and cooperative agreement applications selected for award?
- 700.41 How are contract proposals selected for award?

AUTHORITY: 20 U.S.C. 6011(i).

Subpart A—General

§700.1 What is the purpose of these standards?

- (a) The standards in this part implement section 912(i) of the Educational Research, Development, Dissemination, and Improvement Act of 1994.
- (b) These standards are intended to ensure that activities carried out by the Office of Educational Research and Improvement (the Office) meet the highest standards of professional excellence.

(Authority: 20 U.S.C. 6011(i)(1))

§700.2 What activities must be governed by these standards?

- (a) The standards in this part are binding on all activities carried out by the Office using funds appropriated under section 912(m) of the Educational Research, Development, Dissemination, and Improvement Act of 1994.
- (b) Activities carried out with funds appropriated under section 912(m) of the Act include activities carried out by the following entities or programs:
 - (1) The National Research Institutes.
 - (2) The Office of Reform Assistance and Dissemination.
 - (3) The Educational Resources Information Center Clearinghouses.
 - (4) The Regional Educational Laboratories.
 - (5) The Teacher Research Dissemination Demonstration Program.
 - (6) The Goals 2000 Community Partnerships Program.
 - (7) The National Educational Research Policy and Priorities Board.

(Authority: 20 U.S.C. 6011(i)(1))

§700.3 What additional activities may be governed by these standards?

- (a) The Secretary may elect to apply the standards in this part to activities carried out by the Department using funds appropriated under an authority other than section 912(m) of the Act.
- (b)(1) If the Secretary elects to apply these standards to a competition for new grant or cooperative agreement awards, the Secretary announces, in a notice published in the *FEDERAL REGISTER*, the extent to which these standards are applicable to the competition.
- (2) If the Secretary elects to apply these standards to a solicitation for a contract award, the Secretary announces in the request for proposals the extent to which these standards are applicable to the solicitation.

(Authority: 20 U.S.C. 6011(i))

§700.4 **What definitions apply?**

- (a) Definitions in the Educational Research, Development, Dissemination, and Improvement Act of 1994. The following terms used in this part are defined in 20 U.S.C. 6011(l):
- Development
 - Dissemination
 - Educational Research
 - Office
 - National Research Institute
 - Technical Assistance
- (b) Definitions in Education Department General Administrative Regulations. The following terms used in this part are defined in 34 CFR 77.1:
- Applicant
 - Application
 - Award
 - Department
 - Grant
 - Project
 - Secretary
- (c) Definitions in the Federal Acquisition Regulation. The following terms used in this part are defined in 48 CFR Chapter 1:
- Contracting Officer
 - Employee of an Agency
 - Proposal
 - Solicitation
- (d) Other definitions. The following definitions also apply to this part:
- Act means the Educational Research, Development, Dissemination, and Improvement Act of 1994 (Title IX of Pub.L. 103-227, 108 Stat. 212).
 - EDAR means the Education Department Acquisition Regulation, 48 CFR Chapter 34.
 - EDGAR means the Education Department General Administrative Regulations, 34 CFR Parts 74, 75, 77, 79, 80, 81, 82, 85 and 86.
 - FAR means the Federal Acquisition Regulation, 48 CFR Chapter 1.

(Authority: 20 U.S.C. 6011)

§700.5 What are the processes of open competition?

The Secretary uses a process of open competition in awarding or entering into all grants, cooperative agreements, and contracts governed by these standards. The processes of open competition are the following:

- (a) For all new awards for grants and cooperative agreements, the Secretary will make awards pursuant to the provisions of EDGAR with the exception of the provisions in 34 CFR 75.100(c)(5), 75.200(b)(3), (b)(5), 75.210, and 75.217(b)(1), (b)(2), (c), and (d); and
- (b) For contracts, the Department will conduct acquisitions pursuant to this part in accordance with the requirements of the Competition in Contracting Act, 41 U.S.C. 253, and the FAR.

(Authority: 20 U.S.C. 6011(i)(2); 41 U.S.C. 253)

Subpart B—Selection of Peer Reviewers

§700.10 When is the peer review process used?

The Secretary uses a peer review process—

- (a) To review and evaluate all applications for grants and cooperative agreements and proposals for those contracts that exceed \$100,000;
- (b) To review and designate exemplary and promising programs in accordance with section 941(d) of the Act; and
- (c) To evaluate and assess the performance of all recipients of grants from and cooperative agreements and contracts with the Office.

(Authority: 20 U.S.C. 6011(i)(2)(B))

§700.11 Who may serve as peer reviewers?

- (a) An individual may serve as a peer reviewer for purposes of reviewing and evaluating applications for new awards for grants and cooperative agreements and contract proposals if the individual—
 - (1) Possesses the following qualifications:
 - (i) Demonstrated expertise, including training and experience, in the subject area of the competition.
 - (ii) In-depth knowledge of policy or practice in the field of education.
 - (iii) In-depth knowledge of theoretical perspectives or methodological approaches in the subject area of the competition; and
 - (2) Does not have a conflict of interest, as determined in accordance with §700.12.
- (b) For each competition for new awards for grants and cooperative agreements—

- (i) Department staff may not serve as peer reviewers except in exceptional circumstances as determined by the Secretary; and
- (ii) The majority of reviewers may be persons not employed by the Federal Government.
- (2) For each review of an unsolicited grant or cooperative agreement application—
 - (i) Department employees may assist the Secretary in making an initial determination under 34 CFR 75.222(b); and
 - (ii) Department employees may not serve as peer reviewers in accordance with 34 CFR 75.222(c).
- (c) To the extent feasible, the Secretary selects peer reviewers for each competition who represent a broad range of perspectives.

(Authority: 20 U.S.C. 6011(i)(2)(B))

§700.12 What constitutes a conflict of interest for grants and cooperative agreements?

- (a) Peer reviewers for grants and cooperative agreements are considered employees of the Department for the purposes of conflicts of interest analysis.
- (b) As employees of the Department, peer reviewers are subject to the provisions of 18 U.S.C. 208, 5 CFR 2635.502, and the Department's policies used to implement those provisions.

(Authority: 20 U.S.C. 6011(i)(2)(B))

§700.13 What constitutes a conflict of interest for contracts.

- (a) Peer reviewers for contract proposals are considered employees of the Department in accordance with FAR, 48 CFR 3.104-4(h)(2).
- (b) As employees of the Department, peer reviewers are subject to the provisions of the FAR, 48 CFR Part 3 Improper Business Practices and Personal Conflict of Interest.

(Authority: 41 U.S.C. 423)

Subpart C—The Peer Review Process

§700.20 How many peer reviewers will be used?

- (a) Each application for a grant or cooperative agreement award must be reviewed and evaluated by at least three peer reviewers except—
 - (1) For those grant and cooperative agreement awards under \$50,000, fewer than three peer reviewers may be used if the Secretary determines that adequate peer review can be obtained using fewer reviewers; and
 - (2) For those grant and cooperative agreement awards of more than \$1,000,000, at least five reviewers must be used.

- (b) Each contract proposal must be read by at least three reviewers unless the contracting officer determines that an adequate peer review can be obtained by using fewer reviewers.
- (c) Before releasing contract proposals to peer reviewers outside the Federal Government, the contracting officer shall comply with FAR, 48 CFR 15.413-2 (f).

(Authority: 20 U.S.C. 6011(i) (2) (B))

§700.21 How are applications for grants and cooperative agreements evaluated?

- (a) Each peer reviewer must be given a number of applications to evaluate.
- (b) Each peer reviewer shall—
 - (1) Independently evaluate each application;
 - (2) Evaluate and rate each application based on the reviewer's assessment of the quality of the application according to the evaluation criteria and the weights assigned to those criteria; and
 - (3) Support the rating for each application with concise written comments based on the reviewer's analysis of the strengths and weaknesses of the application with respect to each of the applicable evaluation criteria.
- (c) (1) Except as provided in paragraph (c)(2) of this section, after each peer reviewer has evaluated and rated each application independently, those reviewers who evaluated a common set of applications are convened to discuss the strengths and weaknesses of those applications. Each reviewer may then independently reevaluate and re-rate an application with appropriate changes made to the written comments.
- (2) Reviewers are not convened to discuss an unsolicited application unless the Secretary determines that discussion of the application's strengths and weaknesses is necessary.
- (d) Following discussion and any reevaluation and re-rating, reviewers shall independently place each application in one of three categories, either "highly recommended for funding," "recommended for funding" or "not recommended for funding."
- (e) After the peer reviewers have evaluated, rated, and made funding recommendations regarding the applications, the Secretary prepares a rank order of the applications based solely on the peer reviewers' ratings.

(Authority: 20 U.S.C. 6011(i)(2)(C))

§700.22 How are proposals for contracts evaluated?

- (a) Each peer reviewer must be given a number of technical proposals to evaluate.
- (b) Each peer reviewer shall—
 - (1) Independently evaluate each technical proposal;

- (2) Evaluate and rate each proposal based on the reviewer's assessment of the quality of the proposal according to the technical evaluation criteria and the importance or weight assigned to those criteria; and
- (3) Support the rating for each proposal with concise written comments based on the reviewer's analysis of the strengths and weaknesses of the proposal with respect to each of the applicable technical evaluation criteria.
- (c) After each peer reviewer has evaluated each proposal independently, those reviewers who evaluated a common set of proposals may be convened to discuss the strengths and weaknesses of those proposals. Each reviewer may then independently reevaluate and re-rate a proposal with appropriate changes made to the written comments.
- (d) Following discussion and any reevaluation and re-rating, reviewers shall rank proposals and advise the contracting officer of each proposal's acceptability for contract award as "acceptable," "capable of being made acceptable without major modifications," or "unacceptable." Reviewers may also submit technical questions to be asked of the offeror regarding the proposal.

(Authority: 20 U.S.C. 6011(i)(2)(C))

Subpart D—Evaluation Criteria

§700.30 What evaluation criteria are used for grants and cooperative agreements?

- (a) Except as provided in paragraph (d) of this section, the Secretary announces the applicable evaluation criteria for each competition and the assigned weights in a notice published in the **FEDERAL REGISTER** or in the application package.
- (b) In determining the evaluation criteria to be used in each grant and cooperative agreement competition, the Secretary selects from among the evaluation criteria in paragraph (e) of this section and may select from among the specific factors listed under each criterion.
- (c) The Secretary assigns relative weights to each selected criterion and factor.
- (d) In determining the evaluation criteria to be used for unsolicited applications, the Secretary selects from among the evaluation criteria in paragraph (e) of this section, and may select from among the specific factors listed under each criterion, the criteria which are most appropriate to evaluate the activities proposed in the application.
- (e) The Secretary establishes the following evaluation criteria:
 - (1) **National significance.**
 - (i) The Secretary considers the national significance of the proposed project.
 - (ii) In determining the national significance of the proposed project, the Secretary may consider one or more of the following factors:
 - (A) The importance of the problem or issue to be addressed.

- (B) The potential contribution of the project to increased knowledge or understanding of educational problems, issues, or effective strategies.
- (C) The scope of the project.
- (D) The potential for generalizing from project findings or results.
- (E) The potential contribution of the project to the development and advancement of theory and knowledge in the field of study.
- (F) Whether the project involves the development or demonstration of creative or innovative strategies that build on, or are alternatives to, existing strategies.
- (G) The nature of the products (such as information, materials, processes, or techniques) likely to result from the project and the potential for their effective use in a variety of other settings.
- (H) The extent and quality of plans for disseminating results in ways that will allow others to use the information.

(2) Quality of the project design.

- (i) The Secretary considers the quality of the design of the proposed project.
- (ii) In determining the quality of the design of the proposed project, the Secretary may consider one or more of the following factors:
 - (A) Whether the goals, objectives, and outcomes to be achieved by the project are clearly specified and measurable.
 - (B) Whether there is a conceptual framework underlying the proposed activities and the quality of that framework.
 - (C) Whether the proposed activities constitute a coherent, sustained program of research and development in the field, including a substantial addition to an ongoing line of inquiry.
 - (D) Whether a specific research design has been proposed, and the quality and appropriateness of that design, including the scientific rigor of the studies involved.
 - (E) The extent to which the research design includes a thorough, high-quality review of the relevant literature, a high-quality plan for research activities, and the use of appropriate theoretical and methodological tools, including those of a variety of disciplines, where appropriate.
 - (F) The quality of the demonstration design and procedures for documenting project activities and results.
 - (G) The extent to which development efforts include iterative testing of products and adequate quality controls.
 - (H) The likelihood that the design of the project will successfully address the intended, demonstrated educational need or needs.

- (I) How well and innovatively the project addresses statutory purposes, requirements, and any priority or priorities announced for the program.
- (J) The quality of the plan for evaluating the functioning and impact of the project, including the objectivity of the evaluation and the extent to which the methods of evaluation are appropriate to the goals, objectives, and outcomes of the project.

(3) Quality and potential contributions of personnel.

- (i) The Secretary considers the quality and potential contributions of personnel for the proposed project.
- (ii) In determining the quality and potential contributions of personnel for the proposed project, the Secretary may consider one or more of the following factors:
 - (A) The qualifications, including training and experience, of the project director or principal investigator.
 - (B) The qualifications, including training and experience, of key project personnel.
 - (C) The qualifications, including training and experience, of proposed consultants or subcontractors.

(4) Adequacy of resources.

- (i) The Secretary considers the adequacy of resources for the proposed project.
- (ii) In determining the adequacy of resources for the proposed project, the Secretary may consider one or more of the following factors:
 - (A) The adequacy of support from the lead applicant organization.
 - (B) The relevance and commitment of each partner in the project to the implementation and success of the project.
 - (C) Whether the budget is adequate to support the project.
 - (D) Whether the costs are reasonable in relation to the objectives, design, and potential significance of the project.
 - (E) The potential for continued support of the project after federal funding ends.

(5) Quality of the management plan.

- (i) The Secretary considers the quality of the management plan of the proposed project.
- (ii) In determining the quality of the management plan of a proposed project, the Secretary may consider one or more of the following factors:
 - (A) The adequacy of the management plan to achieve the objectives of the project, including the specification of staff responsibility, timelines, and benchmarks for accomplishing project tasks.
 - (B) The adequacy of plans for ensuring high-quality products and services.
 - (C) The adequacy of plans for ensuring continuous improvement in the operation of the project.

- (D) Whether time commitments of the project director or principal investigator and other key personnel are appropriate and adequate to meet project objectives.
- (E) How the applicant will ensure that a diversity of perspectives are brought to bear in the operation of the project, including those of parents and teachers, where appropriate.
- (F) How the applicant will ensure that persons who are otherwise eligible to participate in the project are selected without regard to race, color, national origin, gender, age, or disability.
- (G) The adequacy of plans for widespread dissemination of project results and products in ways that will assist others to use the information.

(Approved by the Office of Management and Budget under control number 1850-0723)

(Authority: 20 U.S.C. 6011(i)(2)(D)(ii))

§700.31 What additional evaluation criteria shall be used for grants and cooperative agreements?

In addition to the evaluation criteria established in §700.30(e), the Secretary uses criteria or factors specified in the applicable program statute to evaluate applications for grants and cooperative agreements.

(Authority: 20 U.S.C. 6011(i)(2)(D)(ii))

§700.32 What evaluation criteria shall be used for contracts?

- (a) The evaluation criteria to be considered in the technical evaluation of contract proposals are contained in the FAR at 48 CFR 15.605. The evaluation criteria that apply to an acquisition and the relative importance of those factors are within the broad discretion of agency acquisition officials.
- (b) At a minimum, the evaluation criteria to be considered must include cost or price and quality. Evaluation factors related to quality are called technical evaluation criteria.
- (c) Technical evaluation criteria may include, but are not limited to, the following:
 - (1) Technical excellence.
 - (2) Management capability.
 - (3) Personnel qualifications.
 - (4) Prior experience.
 - (5) Past performance.
 - (6) Schedule compliance.

(Authority: 20 U.S.C. 6011(i)(2)(D)(ii))

Subpart E—Selection for Award

§700.40 How are grant and cooperative agreement applications selected for award?

- (a) The Secretary determines the order in which applications will be selected for grants and cooperative agreement awards. The Secretary considers the following in making these determinations:
 - (1) An applicant's ranking.
 - (2) Recommendations of the peer reviewers with regard to funding or not funding.
 - (3) Information concerning an applicant's performance and use of funds under a previous Federal award.
 - (4) Amount of funds available for the competition.
 - (5) Any other information relevant to a priority or other statutory or regulatory requirement applicable to the selection of applications for new awards.
- (b) In the case of unsolicited applications, the Secretary uses the procedures in EDGAR (34 CFR 75.222(d) and (e)).

(Authority: 20 U.S.C. 6022(i)(2)(D)(i))

§700.41 How are contract proposals selected for award?

Following evaluation of the proposals, the contracting officer shall select for award the offeror whose proposal is most advantageous to the Government considering cost or price and the other factors included in the solicitation.

(Authority: 20 U.S.C. 6011(i)(2)(D)(i))

Appendix B
Technical Review Form

TECHNICAL REVIEW FORM
FIELD-INITIATED STUDIES RESEARCH GRANT PROGRAM
[NATIONAL INSTITUTE ON THE EDUCATION OF AT-RISK STUDENTS]
OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT
(CFDA: 84.306F)--FY 1997

REVIEWER SIGNATURE PAGE

APPLICANT ORGANIZATION _____

APPLICATION # 306F7 _____

Reviewer Name: _____

Reviewer Signature: _____ DATE: _____

TECHNICAL REVIEW FORM
FIELD-INITIATED STUDIES RESEARCH GRANT PROGRAM
[NATIONAL INSTITUTE ON THE EDUCATION OF AT-RISK STUDENTS]
OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT
(CFDA: 84.306F)--FY 1997

INITIAL SCORESHEET - TO BE COMPLETED PRIOR TO REVIEW PANEL MEETING

APPLICANT ORGANIZATION _____

APPLICATION # 306F7 _____

<u>SELECTION CRITERIA</u>	<u>MAXIMUM POINTS</u>	<u>POINTS</u>
1. National Significance	30	_____
2. Quality of Project Design	30	_____
3. Quality and Potential Contributions of Personnel	15	_____
4. Adequacy of Resources	15	_____
5. Quality of Management Plan	10	_____
TOTAL of Preliminary Scores	100	_____

Based on an overall assessment of this application, please check your recommendation:

_____ Highly Recommended for Funding

_____ Recommended for Funding

_____ Not Recommended for Funding

TECHNICAL REVIEW FORM
FIELD-INITIATED STUDIES RESEARCH GRANT PROGRAM
[NATIONAL INSTITUTE ON THE EDUCATION OF AT-RISK STUDENTS]
OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT
(CFDA: 84.306F)--FY 1997

APPLICANT ORGANIZATION _____

APPLICATION # 306F7 _____

<u>SELECTION CRITERIA</u>	<u>MAXIMUM POINTS</u>	<u>POINTS</u>
1. National Significance	30	_____
2. Quality of Project Design	30	_____
3. Quality and Potential Contributions of Personnel	15	_____
4. Adequacy of Resources	15	_____
5. Quality of Management Plan	10	_____
TOTAL	100	_____

Based on an overall assessment of this application, please check your recommendation:

_____ Highly Recommended for Funding

_____ Recommended for Funding

_____ Not Recommended for Funding

1. National Significance. (30 points)

The Secretary considers the national significance of the proposed project.

In determining the national significance of the proposed project, the

Secretary considers the following factors—

(A) The importance of the problem or issue to be addressed.

(B) The potential contribution of the project to increased knowledge or understanding of educational problems, issues, or effective strategies.

(C) The potential contribution of the project to the development and advancement of theory and knowledge in the field of study.

STRENGTHS:

Maximum points 30

(excellent) 26-30

(good) 21-25

(fair) 16-20

(poor) 1-15

(missing) 0

SCORE _____

WEAKNESSES:

2. Quality of Project Design (30 points)

The Secretary considers the quality of the design of the proposed project. In determining the quality of the design of the proposed project, the Secretary considers the following factors--

(A) Whether the goals, objectives, and outcomes to be achieved by the project are clearly specified and measurable.

(B) Whether a specific research design has been proposed and the quality and appropriateness of that design, including the scientific rigor of the studies involved.

STRENGTHS:Maximum points 30

(excellent) 26-30

(good) 21-25

(fair) 16-20

(poor) 1-15

(missing) 0

SCORE _____

WEAKNESSES:

BEST COPY AVAILABLE

3. Quality and potential contributions of personnel (20 points)

The Secretary consider the quality and potential contributions of the personnel for the proposed project. In determining the quality and potential contributions of the personnel for the the proposed project, the Secretary considers the following factors—

(A) The qualifications, including training and experience, of the project director or principal investigator.

(B) The qualifications, including training and experience, of key project personnel.

STRENGTHS: .

Maximum points 15

(excellent) 14-15

(good) 11-13

(fair) 8-10

(poor) 1-7

(missing) 0

SCORE —

WEAKNESSES:

4. Adequacy of Resources (10 points)

The Secretary considers the adequacy of the resources for the proposed project. In determining adequacy of resources for the proposed project the Secretary considers the following factors--

(A) Whether the budget is adequate to support the project.

(B) Whether the costs are reasonable in relation to the objectives, design, and potential significance of the project.

STRENGTHS:

Maximum points 15

(excellent) 14-15

(good) 11-13

(fair) 8-10

(poor) 1-7

(missing) 0

SCORE _____

WEAKNESSES:

BEST COPY AVAILABLE

5. Quality of the Management Plan (10 points)

The Secretary considers the quality of the management plan of the proposed project. In determining the quality of the management plan of a proposed project, the Secretary considers the following factors—

(A) The adequacy of the management plan to achieve the objectives of the project, including the specification of staff responsibility, timelines, and benchmarks for accomplishing project tasks.

(B) Whether time commitments of the project director or principal investigator and other key personnel are appropriate and adequate to meet project objectives.

(C) How the applicant will ensure persons who are otherwise eligible to participate in the project are selected without regard to race, color, national origin, gender, age or disability.

STRENGTHS:

Maximum points 10

(excellent) 9-10

(good) 7-8

(fair) 6-7

(poor) 1-5

(missing) 0

SCORE

WEAKNESSES:

SUMMARY SHEET

OVERALL COMMENTS: Comments should support your recommendation.

Please describe strengths and weaknesses. Include suggestions to improve the project in future submissions

Appendix C

Evaluation Criteria for Peer Review

Appendix C

Evaluation Criteria for Peer Review

For the fiscal year 1997 FIS competition, the broad evaluation criteria, the points assigned to each, and the associated specific factors are as follows:

1. National Significance (30 points)

Specific factors: (a) the importance of the problem or issue to be addressed; (b) the potential contribution of the project to increased knowledge or understanding of educational problems, issues, or effective strategies; and (c) the potential contribution of the project to the development and advancement of theory and knowledge in the field of study.

2. Quality of Project Design (30 points)

Specific factors: (a) whether the goals, objectives, and outcomes to be achieved by the project are clearly specified and measurable; and (b) whether a specific research design has been proposed, and the quality and appropriateness of that design, including the scientific rigor of the studies involved.

3. Quality and Potential Contribution of Personnel (20 points)

Specific factors: (a) the qualifications, including training and experience, of the project director or principal investigator; and (b) the qualifications, including training and

experience, of key project personnel.

4. Adequacy of Resources (10 points)

Specific factors: (a) whether the budget is adequate to support the project; and
(b) whether the costs are reasonable in relation to the objectives, design, and potential significance of the project.

5. Quality of Management Plan (10 points)

Specific factors: (a) the adequacy of the management plan to achieve the objectives of the project, including the specification of staff responsibility, timelines, and benchmarks for accomplishing the project tasks; (b) whether time commitments of the project director or principal investigator and other key personnel are appropriate and adequate to meet project objectives; and (c) how the applicant will ensure that persons who are otherwise eligible to participate in the project are selected without regard to race, color, national origin, gender, age, or disability.

For the fiscal year 1996 FIS competition, the point structure was different, with 15 points each being assigned to quality of personnel and adequacy of resources.¹ For center competitions in both 1996 and 1997, the broad criteria were the same as for FIS, but the specific factors were somewhat different:

¹ Some FIS 1996 review forms show only the broad criteria, not the specific factors.

1. National Significance (30 points)

Specific factors: (a) the importance of the problem or issue to be addressed; (b) the potential contribution of the project to increased knowledge or understanding of educational problems, issues, or effective practice; (c) the potential contribution of the project to the development and advancement of theory and knowledge in the field of study; and (d) the nature of the products (such as information, materials, processes, or techniques) likely to result from the project and the potential for their effective use in a variety of other settings.

2. Quality of Project Design (30 points)

Specific factors: (a) whether there is a conceptual framework underlying the proposed activities and the quality of that framework; (b) whether the proposed activities constitute a coherent, sustained program of research and development in the field, including a substantial addition to an ongoing line of inquiry; (c) the extent to which the research design includes a thorough, high-quality review of the relevant literature, a high-quality plan for research activities, and the use of appropriate theoretical and methodological tools, including those of a variety of disciplines, where appropriate; and (d) the quality of the plan for evaluating the function and impact of the project, including the objectivity of the evaluation and the extent to which the methods of evaluation are appropriate to the goals, objectives, and outcomes of the project.

3. Quality and Potential Contribution of Personnel (20 points)

Specific factors: (a) the qualifications, including training and experience, of the project director or principal investigator; (b) the qualifications, including training and experience, of key project personnel; and (c) whether the applicant has assembled a group of high-quality researchers sufficient to achieve the mission of the center.

4. Adequacy of Resources (10 points)

Specific factors: (a) the adequacy of support from the lead applicant organization; (b) the relevance and commitment of each partner in the project to its implementation and success; (c) whether the costs are reasonable in relation to the objectives, design, and potential significance of the project; (d) whether the proposed organizational structure and arrangements will facilitate achievement of the mission of the center; and (e) whether the directors and support staff will devote a majority of their time to the activities of the center.

5. Quality of Management Plan (10 points)

Specific factors: (a) the adequacy of the management plan to achieve the objectives of the project, including the specification of staff responsibility, timelines, and benchmarks for accomplishing the project tasks; (b) the adequacy of plans for ensuring high-quality products and services; (c) how the applicant will ensure that a diversity of perspectives are brought to bear in the operation of the project, including those of parents and teachers, where appropriate; (d) whether there is substantial staff commitment to the work of the center; (e) the contributions of primary researchers (other than researchers at the proposed

center), the appropriateness of such researchers' experience and expertise in the context of the proposed center activities, and the adequacy of such primary researchers' time and commitment to the achievement of the mission of the center; and (f) the manner in which the results of education research will be disseminated for further use, including how the center will work with the Office of Reform Assistance and Dissemination.



PPB 1999-6314g